EXHIBIT I

FALS

TRENDS in Biochemical Sciences TIBS

The official publication of the INTERNATIONAL UNION OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

Camelid antibodies

GEFs

Bacterial chemoreceptors

Prion protein interactions

Recognition of antigens by single-domain antibody fragments: the superfluous luxury of paired domains

Serge Muyldermans, Christian Cambillau and Lode Wyns

The antigen-binding site of antibodies from vertebrates is formed by combining the variable domains of a heavy chain (VH) and a light chain (VL). However, antibodies from camels and llamas are an important exception to this in that their sera contain, in addition, a unique kind of antibody that is formed by heavy chains only. The antigen-binding site of these antibodies consists of one single domain, referred to as VHH. This article reviews the mutations and structural adaptations that have taken place to reshape a VH of a VH–VL pair into a single-domain VHH with retention of a sufficient variability. The VHH has a potent antigen-binding capacity and provides the advantage of interacting with novel epitopes that are inaccessible to conventional VH–VL pairs.

Serge Muyldermans*
Lode Wyns
Vrije Universiteit Brussel, Vlaams Interuniversitair Instituut Biotechnologie, Paardenstraat 65, B-1640 Sint Genesius Rode, Belgium.
*e-mail: svmuylde@vub.ac.be

Christian Cambillau
Architecture et Fonction des Macromolecules Biologiques, UMR 6098, CNRS and Université de la Méditerranée, 31 Chemin Joseph Aiguier, 13402 Marseille Cedex 20, France.

Probably the fastest and most efficient protein-engineering task is performed daily by the immune system of vertebrates. The immune system produces large amounts of highly specific adapter molecules – known as immunoglobulins or antibodies – that are raised to virtually all possible foreign molecules, either small or large.

The antigen-binding site of antibodies

The efficient design of such 'emergency' molecules relies on the possibility of generating an enormous number of different antibody molecules as B-cell receptors. The B cells carrying an antibody that recognizes the antigen through surface and charge complementarity will proliferate and evolve into cells that produce soluble antibodies. Despite having the capacity to produce a nearly unlimited variation in antigen-binding surfaces, the antibody molecules have a composition and overall structure that is remarkably well conserved throughout the vertebrate phylum.

Antibodies are formed by two identical heavy and two identical light polypeptide chains, folded in four and two domains, respectively[1] (Fig. 1a). The N-terminal domain of each chain is more variable in sequence than the others. Sequence variations in the variable domain of the heavy and light chain (VH and VL, respectively) can be introduced at any of the multiple stages of antibody formation[2] (Box 1). Sequence comparison of the various VH or VL domains indicated that there are six regions [three in VH (Fig. 2) and three in VL] in which the amino acid sequence is more variable than the remainder of the sequence. It was immediately hypothesized that these hypervariable regions would interact with the antigen, and they were therefore named CDRs, for complementarity determining regions. (Within each domain, the CDRs are numbered 1–3 following their occurrence in sequence.) Indeed, crystallographic data confirmed that the folded VH and VL domains associate so that the six hypervariable loops juxtapose at one end of the molecule to form a continuous surface of ~1000 Å^2. Antibody–antigen complexes showed that 190–350 Å^2 or 400–900 Å^2 of this surface actually interacts with haptens or proteinaceous antigens, respectively[3].

All CDRs can potentially make contact with the antigen, although very seldom all at the same time for anti-hapten or anti-peptide antibodies. The antigen contacts made by CDR3 are generally more extensive. These C-terminally located hypervariable regions are the most variable in sequence and length[4] because they contain amino acids encoded by codons that are generated by V(D)J recombination, including its imprecise joinings [see Box 1 for V(D)J recombination]. An 'antigen-eye view' of the antigen-binding site reveals that the highest variability is at the centre of the site, where the CDR3 loops of VH and VL cluster. Most likely, these loops will dictate the antigen specificity. The somatic mutations introduced at a later stage, during the affinity maturation towards a specific antigen, seem to occur preferentially at the periphery of the antigen-binding site[5]. The size of the third hypervariable loop of VH, in conjunction with the flexible association of VH and VL at various angles and distances, generates a structural diversity of binding sites that can be grouped into three major classes. These are schematically described as cavities, grooves and planar sites, and they correspond to the size and type (hapten, peptide and protein, respectively) of the antigen[6].

Canonical loop structures

At a structural level, the CDRs fold into one of a limited number of so-called canonical structures[7]. On the basis of the sequence similarity of key sites in the CDR, it was proposed that the same canonical structures would also shape the antigen-binding site of the cartilaginous fish antibodies[8]. These fish are the most distantly related species to human (of those species that are known to have an immune system), which indicates that the canonical loop structures arose early in the evolution of the immune system and

http://tibs.trends.com 0968-0004/01/$ – see front matter PII: S0968-0004(01)01790-X

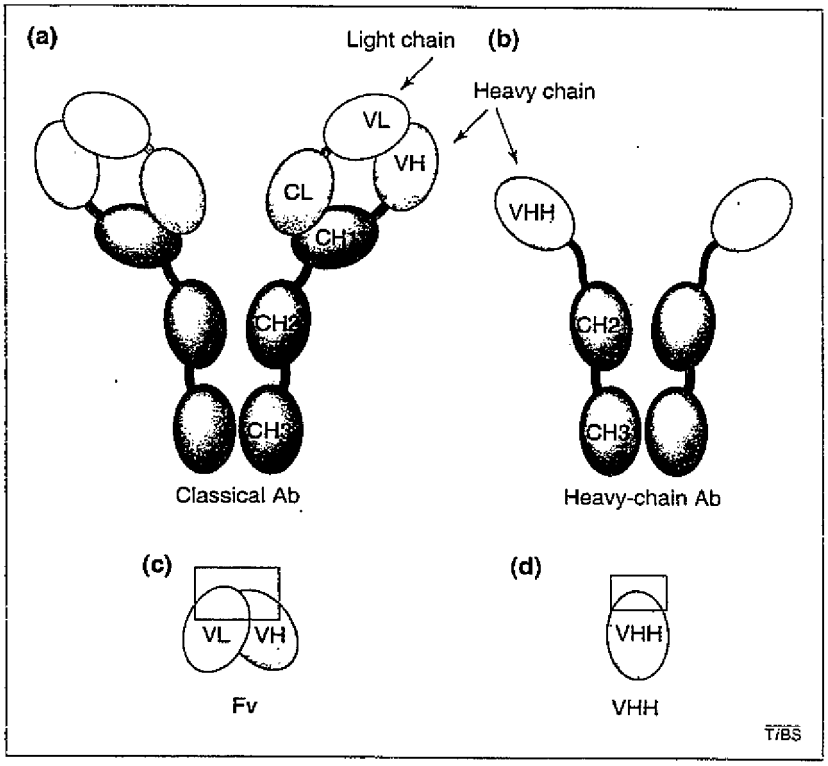


Fig. 1. Schematic representation of (a) a classical antibody, (b) a heavy-chain antibody of camelids, and (c and d) their respective antigen-binding fragments. The antigen-binding site of the VH–VL pair or of the VHH is denoted by the red boxes. The yellow bars in b and d indicate a frequently occurring inter-loop disulphide bond within the VHH. Abbreviations: Ab, antibody; CH1–3, first, second and third constant heavy-chain domains; CL, constant light-chain domain; Fv, variable fragment; VH, variable domain of heavy chain; VHH, variable domain of heavy chain of heavy-chain antibody; VL, variable domain of light chain.

were conserved thereafter. The canonical loop structures were defined originally by the Cα positions, but later analysis also considered the peptide backbone and this led to the subdivision of some canonical structures into subtypes[9]. The comparison between the sequence and crystallographic structures resulted in algorithms that allow the prediction of the conformation of five out of the six hypervariable loops[10,11]; the CDR3 of VH, the most variable CDR in length and amino acid composition, is more difficult to predict. New crystallographic data has provided novel insights into this loop structure[12,13], but it remains inadvisable to make sweeping generalizations regarding CDR3 loop conformations given that there are ~10^9 different antibody specificities of which only very few complexes have been investigated so far[3].

Heavy-chain antibodies

In the natural world, the generation of functional antibody genes and the final composition of antibodies are seemingly so adequate that it is difficult to imagine improvements to, or even deviations from, the natural state. But surprisingly, the Tylopoda (camels, dromedaries and llamas) have developed an additional antibody molecule with a homodimeric heavy-chain composition that is devoid of light chains[14]; such immunoglobulins are called 'heavy chain antibodies' (Fig. 1b). Consequently, the antigen-binding fragment of these heavy-chain antibodies is confined to one single domain (i.e. the variable domain referred to as VHH for variable domain of the heavy chain of a heavy-chain antibody) (Fig. 1d), instead of the paired VH and VL domains (Fig. 1c). The immunization of camelids showed that the response in conventional or heavy-chain IgG depended on the type of antigen[15].

VH and VHH sequence difference

Obviously, the amino acid sequences of the variable domain of the naturally occurring heavy-chain antibodies would be expected to acquire important adaptations to compensate for the absence of association with the light-chain variable domain. Nevertheless, the VH and VHH amino acid sequences

Box 1. Major mechanisms to build up the diversity of antigen-binding sites

A quasi-infinite number of different antibodies can be generated, each having a unique antigen-binding site formed by the N-terminal domain of the heavy and light chains[a]. Several distinct mechanisms are at the origin of this antigen-binding site diversity. First, a recombinatorial diversity is obtained by random selection of one variable heavy-chain gene (VH), one diversity gene (D) and one heavy joining (JH) minigene, or one variable light-chain gene (VL) and one light joining (JL) gene segment out of a pool, to constitute the VH and VL domains, respectively. Second, a junctional diversity is added by the imprecise joining mechanisms and by deletion or addition of random nucleotides at the borders of the recombining VH–D–JH minigenes. Third, a combinatorial diversity created by the assembly of the VH and the VL domain completes the antigen-binding site. Fourth, the architecture of the antigen-binding site is enlarged by adjusting the angle between the associated VL and VH domains[a].

The primary antigen-binding site then benefits from a specific maturation event by the acquisition of somatic hypermutations that improve the shape complementarity of the antibody with the target antigen. The underlying mechanism of the somatic hypermutation is still unclear; however, it is well established that it is dictated by DNA hotspots such as AGY and TAY (Ref. b). The human immune system is also harnessed with a powerful selection mechanism so that only the cells producing the best binders survive.

References
a Padlan, E.A. (1994) Anatomy of the antibody molecule. *Mol. Immunol.* 31, 169–217
b Milstein, C. *et al.* (1998) Both DNA strands of antibody genes are hypermutation targets. *Proc. Natl. Acad. Sci. U. S. A.* 95, 8791–8794

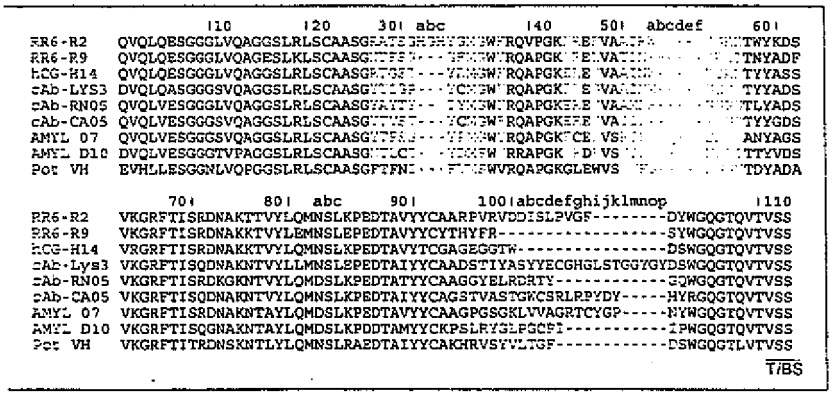


Fig. 2. Alignment of VHH sequences with known crystallographic structure and of one human variable domain of heavy chain (VH) (Pot VH). The hypervariable regions are shown in red, green and blue. The VHH hallmark amino acids are in pink, and the Cys residues involved in either an intradomain (C22 and C92) or an inter-loop disulphide bond are highlighted in yellow. The two sequences at the top are against a hapten (azo-dye Reactive Red, RR6); all other VHHs are directed against proteins. Abbreviations: AMYL, amylase; CA, carbonic anhydrase; hCG, human chorionic gonadotropin hormone; LYS, lysozyme; RN, RNase A.

share a high degree of identity (Fig. 2) and are most similar (~80%) to the human VH of family III (Ref. 16), the most common human VH family in terms of both expression and genome complexity[17]. The amino acids at positions that determine the typical immunoglobulin fold[18] are all well conserved in the VHH. However, four amino acids that are extremely well conserved in all VHs are constitutively substituted in the VHH. These residues [Val37Phe (or Tyr), Gly44Glu (or Gln), Leu45Arg (or Cys) and Trp47Gly (or Ser, Leu, Phe) (Fig. 2)] discriminate the conventional VH from the heavy-chain specific VHH.

Three hypervariable regions can be clearly distinguished in the VHH sequences, although the average variability of the remaining parts is increased relative to that in human or mouse VH (Ref. 19). In addition, the CDR3 is longer in VHH, on average, than in VHs (17, 12 and 9 amino acids in dromedary VHHs, human VH and mouse VH, respectively[16]).

Dromedary HCAbs are generated from a limited number of diverse V_HH germline segments

The unique, functional, heavy-chain IgG antibodies occur (to the best of our knowledge) exclusively within the Tylopoda. It is expected that their appearance must be paralleled by gene adaptations, and altered gene organization and usage. Indeed, it seems that new and dedicated sets of immunoglobulin genes arose in the common ancestor of the camelids. Separate VH and VHH germline genes, probably residing within the same locus, recombine with common D and JH gene segments to form a VH or VHH domain, respectively[19]. A limited number of VHH germline genes (~40) have been identified in the dromedary genome: this number is approximately half that of functional human VH genes[20]. However, the repertoire of the primary VHH domains is apparently further diversified by active somatic mutation mechanisms[19].

The dromedary VHH germline genes encode a Cys residue in CDR1 (or in the framework region at position 45). Cys residues at these positions are absent in all VH germline genes including those of the dromedary[19]. In addition, a second Cys, in the VHH CDR3, is introduced exclusively during the recombination of the VHH–D–JH genes. These additional Cys residues form an inter-loop disulphide bond that stabilizes the VHH domain[21]. Furthermore, this bond is expected to impose conformational restraints on the loop flexibility in the absence of antigen so that the entropic penalty upon antigen binding is minimized.

VHH structure

Polymerase chain reaction and phage display are routine techniques used to clone the antigen-binding modules (VH–VL pairs or VHHs; Figs 1c,d) from antibodies and to select antigen-specific binders[22]. It has been shown repeatedly that the selected VHH fragments can be expressed extremely well as soluble proteins in bacteria and yeast[23]. Several of these recombinant VHHs directed against haptens or various proteins were crystallized with or without their antigen. This structural information confirmed

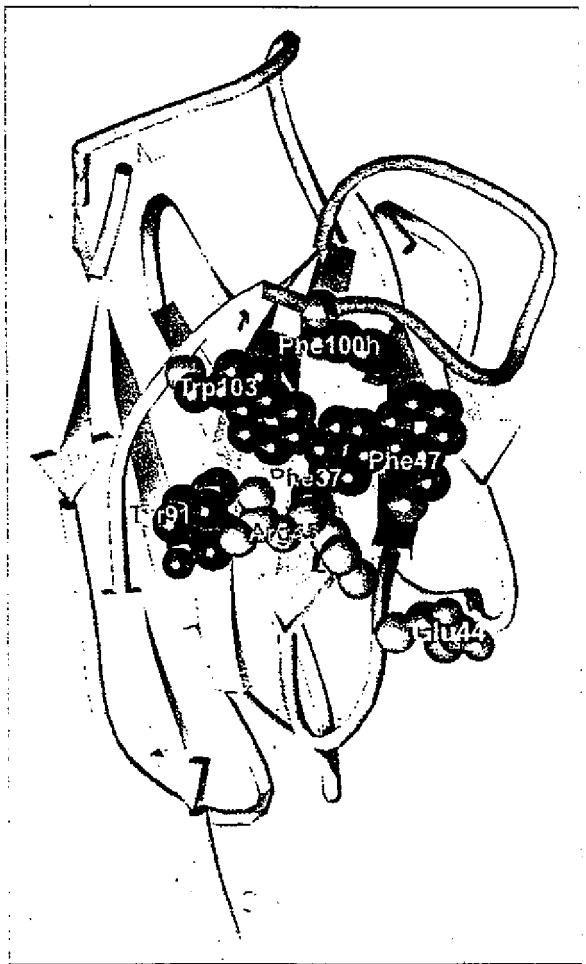


Fig. 3. The immunoglobulin fold of a VHH (R2-VHH)[27]. The scaffold is in yellow (arrows are β strands), and the CDRs 1, 2 and 3 are in red, green and blue, respectively. The hydrophobic cluster of Phe37, Phe47, Tyr91, Trp103 and the Phe100h (g on fig) at three amino acids upstream from Trp103 are shown in purple. The VHH hallmark amino acids Arg45 and Glu44 at the 'VL-side' of the VHH domain are shown in cyan and red, respectively.

that the VHH scaffold adopts the typical immunoglobulin fold (Fig. 3) and superimposes perfectly on the conventional VH structure[24–27]. Furthermore, the crystal structures clarify the necessary adaptations that took place in the VHH domain to cope with the absence of a VL domain. These adaptations are concentrated in two areas: the 'side' of the domain that corresponds to the VH-side contacting the VL, and the CDRs themselves.

'VL side'

The amino acids that constitute the VH–VL interface consist of both conserved residues and hypervariable residues[1]. It is suggested that the latter modulate the relative position of the VH and VL, potentially altering the specificity of the antibody. Removing the VL domain exposes a large hydrophobic patch of the VH to the aqueous solvent. The side chains of Val37, Gln39, Gly44, Leu45, Trp47, Tyr91 and Trp103, residues of the conserved framework, have fixed positions in all VH–VL pairs and provide an interacting surface of ~700 Å^2 (Ref. 1). As expected, the exposure of such a large hydrophobic region to solvent leads to aggregation or stickiness of an isolated VH domain[28]. The VHH-specific amino acid substitutions cluster in this region and render the area much more hydrophilic than it would be otherwise. This hydrophilicity is augmented further by rotation of the side chains of adjacent residues without deforming the Cα backbone. For example, the Trp103 side chain rotates over its Cβ–Cγ bond to expose its most polar part, the Nε, to the environment. Phe37 fills a hydrophobic pocket created by the side chains of Phe47, Tyr91, Trp103 and the CDR3, where the Phe residue three amino acids upstream of Trp103 plays a central role (Fig. 3). In addition, the VHH CDR3 folds over this part of the domain, and covers some of the amino acids that are buried by the VL partner in a typical VH–VL dimer. This reshaped surface explains both the failure of a VHH to associate with a VL domain and the increased solubility of an isolated VHH domain[29].

Riechmann described the nuclear magnetic resonance structure of a partial 'camelised' human VH (i.e. a few amino acids were substituted into framework two of a human VH to mimic a camel VHH; Ref. 30). Mutation of the Leu45 into Arg (and Trp47 into Ile to improve expression levels) rendered the isolated human VH domain more soluble. However, these mutations induced unexpected backbone deformations at positions 37–38 and 45–47, and the side chain of Trp103 took a completely new position. Thus, it appears that the backbone scaffold of the original VH and VHH are superimposable, whereas the partial 'camelisation' of a human VH by Ile47Gly and Val37Phe mutations introduces bulges and side-chain deformations at adjacent β strands.

Antigen-binding loops

In the absence of the VH–VL combinatorial diversity, new mechanisms have to be introduced to increase the diversity of the antigen-binding loops within one domain. In addition, there needs to be compensation for the loss of antigen-interacting surface contributed by the hypervariable loops of the VL. Apparently, this has been achieved largely because of distinct differences in the organization of the hypervariable loops. First, the CDR1 of VHHs is extended towards the N-terminal end (Fig. 2). These amino acids form a loop connecting two β strands of adjacent sheets in the immunoglobulin domain. In VHHs this region is more variable in sequence than in VHs, probably because of the acquisition of somatic mutations that are selected during the affinity maturation process. Indeed, DNA mutational hotspots (i.e. DNA sequences that are more susceptible to mutation[31]) are imprinted in this region of the VHH germline genes but not in VHs (Ref. 19), and the amino acids encoded by these hotspot sequences were proven to interact directly with antigen[24,25]. Second, the conformation of the CDR1 and CDR2 in a VHH often deviates from canonical structures in human or mouse VHs (Ref. 32). However, there is no *a priori* reason why the CDR1 and CDR2 of conventional VHs could not adopt these alternative conformations. Third, the CDR3 of a VHH is, on average, longer than that of VHs, and is also more accessible to solvent, thus creating a larger surface area available for antigen interaction.

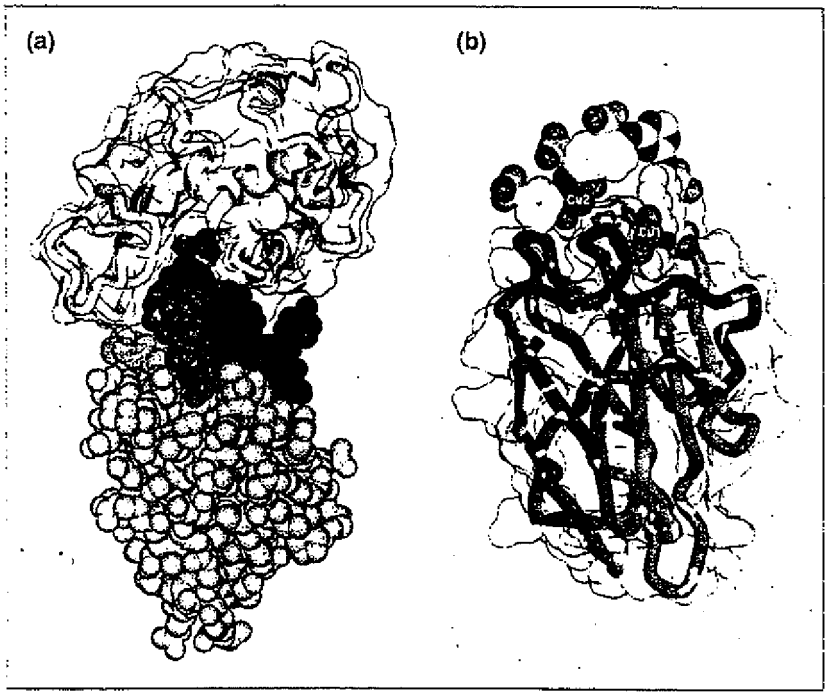


Fig. 4. (a) The structure of the single-domain VHH in complex with lysozyme (yellow). The VHH is shown with its complementarity determining region (CDR)1, CDR2 and CDR3 in red, green and blue respectively. The shape complementarity between the convex paratope and the active site of lysozyme is striking. (b) A Cα representation of the structure of the azo-dye Reactive-Red (RR6)-binder in complex with its hapten (space fillings). The cavity between the CDR1 (red) and CDR2 (green) provides a pocket for the hapten. The CDR3 (blue) serves mainly to cover the 'VL-side' of the VHH domain. The presence of the two copper ions, Cu1 and Cu2, is indicated.

VHH-antigen binding

The presence of an enlarged CDR1 and CDR3 in VHHs, and loops that exhibit alternative canonical

structures, increases the structural repertoire of the antigen-binding site in the single-domain VHH and compensates for the absence of the three VL CDRs. Altogether, the architecture of the antigen-binding site (paratope), and the antigen interaction mode, of VHHs are very diverse. As expected, planar paratopes are observed to interact with a proteinaceous antigen. But surprisingly, in an RNase A-binder, the contacts are only made with two loops[24], CDR1 and CDR3, whereas in the carbonic anhydrase-binder only one single loop, CDR3, is involved in antigen interaction. In another example, ten consecutive amino acids of the CDR3 protrude from the antigen-binding site and penetrate into the active site of lysozyme where they mimic the natural substrate of the enzyme[33] (Fig. 4a). Such a striking formation of a large convex paratope by a protruding CDR loop has not yet been observed in conventional antibodies. Furthermore, it provides an antigen–VHH interface area of ~1700 Å^2, which is as large as the interface between antigens and a VH–VL pair[34]. Finally, a third paratope architecture was observed for a hapten binder. Haptens are normally captured in a groove or cavity at the VH–VL interface[3,6]. Despite the absence of the VL, a llama VHH was able to form a cavity with its three CDRs to accommodate the hapten[27] (Fig. 4b); however, the importance of CDR3 in the interaction was reduced.

Concluding remarks

The recombinant VHH is a minimal-sized, intact antigen-binding domain derived from *in vivo* matured camel or llama heavy-chain antibodies. It is extremely stable and binds antigen with affinities in the nanomolar range. The absence of a VL domain permits the VHHs many structural variations that are not permissible in VH domains associated with VLs. Some of the amino acids within the VH CDRs participate in the VH–VL contact, and mutations at these spots are consequently forbidden or at least limited. Such a restraint is not an issue for the amino acids within the CDRs of VHHs. The subtle substitutions and structural rearrangements that occur between the VL-interacting surface of a VH and the corresponding side of the VHH provide an example of natural protein engineering converting a heterodimeric protein into a monomeric protein.

The absence of VL-antigen contacts is compensated for in a VHH by an extended CDR1 and a more exposed CDR3. The structural repertoire of the antigen-binding site of VHHs is therefore diverse, and new canonical structures have been identified. Besides the standard architectures such as cavities and planar surfaces, the antigen-binding site of VHH also includes protruding loops. Thus, a monomeric domain that interacts with antigens also has advantages. Indeed, the absence of the VL domain means that the paratope is concentrated over a smaller area so that small, hidden epitopes can still be targeted. It seems that heavy-chain antibodies or conventional antibodies recognize different antigenic sites. For example, in contrast to conventional antibodies, the camel heavy-chain antibodies interact preferentially with the active site of enzymes[35]. Hence the VHHs directed against enzymes often appear to be potent inhibitors. The design of small enzyme inhibitors derived from the antigen-binding site of the VHHs will be simplified owing to the reduced complexity of the VHH paratope, which will have three instead of six antigen-binding loops. However, it is expected that single-domain antibodies of camelids will find their way into many more biotechnological applications, especially whenever large quantities are needed at low cost, and when immunofusions, immunomodulation or robust molecular recognition units in affinity adsorbents, or in antibody-based protein chips, are envisaged[36–38].

Acknowledgements
We thank our colleagues S. Spinelli, M. Tegoni, A. Desmyter, T. Transue and K. Decanniere for crystallography, M.A. Ghahroudi, M. Lauwereys, L. Frenken and V.K. Nguyen for VHH identification and ontogenic studies, and E. Blanc for help with the figures. The work in our laboratories was supported by 'Biotechnology' EEC programmes and by the ESA grant.

References

1 Padlan, E.A. (1994) Anatomy of the antibody molecule. *Mol. Immunol.* 31, 169–217
2 Tonegawa, S. (1983) Somatic generation of antibody diversity. *Nature* 302, 575–581
3 Padlan, E.A. (1996) X-ray crystallography of antibodies. *Adv. Protein Chem.* 49, 57–133
4 Wu, T.T. *et al.* (1993) Length distribution of CDR H3 in antibodies. *Protein Struct. Funct. Genet.* 16, 1–7
5 Tomlinson, I.M. *et al.* (1996) The imprint of somatic hypermutation on the repertoire of human germline V genes. *J. Mol. Biol.* 256, 813–817
6 Webster, D.M. *et al.* (1994) Antibody–antigen interactions. *Curr. Opin. Struct. Biol.* 4, 123–129
7 Chothia, C. and Lesk, A.M. (1987) Canonical structures for the hypervariable regions of immunoglobulins. *J. Mol. Biol.* 196, 901–917
8 Barré, S. *et al.* (1994) Structural conservation of hypervariable regions in immunoglobulin evolution. *Nat. Struct. Biol.* 1, 915–920
9 Al-Lazikani, B. *et al.* (1997) Standard conformations for the canonical structures of immunoglobulins. *J. Mol. Biol.* 273, 927–948
10 Chothia, C. *et al.* (1989) Conformations of immunoglobulin variable regions. *Nature* 342, 877–883
11 Martin, A.C.R. and Thornton, J.M. (1996) Structural families in loops of homologous proteins: automatic classification, modelling and application to antibodies. *J. Mol. Biol.* 263, 800–815
12 Morea, V. *et al.* (1998) Conformations of the third hypervariable region in the VH domain of immunoglobulins. *J. Mol. Biol.* 275, 269–294
13 Shirai, H. *et al.* (1999) H3-rules: identification of CDR3-H3 structures in antibodies. *FEBS Lett.* 455, 188–197
14 Hamers-Casterman, C. *et al.* (1993) Naturally occurring antibodies devoid of light chains. *Nature* 363, 446–448
15 van der Linden, R. *et al.* (2000) Induction of immune reponses and molecular cloning of the heavy chain antibody repertoire of *Lama glama*. *J. Immunol. Methods* 240, 185–195
16 Vu, K.B. *et al.* (1997) Comparison of llama VH sequences from conventional and heavy chain antibodies. *Mol. Immunol.* 34, 1121–1131
17 Tomlinson, I.M. *et al.* (1992) The repertoire of human germline VH sequences reveals about fifty groups of VH segments with different hypervariable loops. *J. Mol. Biol.* 227, 776–798
18 Chothia C. *et al.* (1998) Structural determinants in the sequences of immunoglobulin variable domain *J. Mol. Biol.* 278, 457–479
19 Nguyen, V.K. *et al.* (2000) Camel heavy-chain antibodies: diverse germline VHH and specific mechanisms enlarge the antigen-binding repertoire. *EMBO J.* 19, 921–931
20 Cook, G.P. and Tomlinson, I.M. (1995) The human immunoglobulin VH repertoire. *Immunol. Today* 16, 237–242
21 Davies, J. and Riechmann, L. (1996) Single antibody domains as small recognition units: design and *in vitro* antigen selection of camelised, human VH domains with improved protein stability. *Protein Eng.* 9, 531–537
22 Hoogenboom, H.R. and Chames, P. (2000) Natural and designer binding sites made by phage display technology. *Immunol. Today* 21, 371–378
23 Frenken, L. *et al.* (2000) Isolation of antigen specific llama VHH antibody fragments and their high level secretion by *S. cerevisiae*. *J. Biotechnol.* 78, 11–21
24 Decanniere, K. *et al.* (1999) A single-domain antibody fragment in complex with RNase A: non-canonical loop structures and nanomolar affinity using two CDR loops. *Structure* 7, 361–370

25 Desmyter, A. *et al.* (1996) Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme. *Nat. Struct. Biol.* 3, 803–811
26 Spinelli, S. *et al.* (1996) The crystal structure of a llama heavy chain variable domain. *Nat. Struct. Biol.* 3, 752–757
27 Spinelli, S. *et al.* (2000) Camelid heavy-chain variable domains provide efficient combining sites to haptens. *Biochemistry* 39, 1217–1222
28 Ward, E.S. *et al.* (1989) Binding activities of a repertoire of single immunoglobulin variable domains secreted from *E. coli*. *Nature* 341, 544–546
29 Davies, J. and Riechmann, L. (1994) Camelising human antibody fragments: NMR studies on VH domains. *FEBS Lett.* 339, 285–290
30 Riechmann, L. (1996) Rearrangement of the former VL interface in the solution structure of a camelised, single domain VH antibody. *J. Mol. Biol.* 259, 957–969
31 Milstein, C. *et al.* (1998) Both DNA strands of antibody genes are hypermutation targets. *Proc. Natl. Acad. Sci. U. S. A.* 95, 8791–8794
32 Decanniere, K. *et al.* (2000) Canonical antigen binding loop structures: more structures, more canonical classes? *J. Mol. Biol.* 300, 83–91
33 Transue, T.R. *et al.* (1998) Camel single domain antibody inhibits enzyme by mimicking carbohydrate substrate. *Protein Struct. Funct. Genet.* 32, 515–522
34 Lo Conte, L. *et al.* (1999) The atomic structure of protein–protein recognition sites. *J. Mol. Biol.* 285, 2177–2198
35 Lauwereys, M. *et al.* (1998) Potent enzyme inhibitors derived from dromedary heavy-chain antibodies. *EMBO J.* 17, 3512–3120
36 Sheriff, S. and Constantine, K.L. (1996) Redefining the minimal antigen-binding fragment. *Nat. Struct. Biol.* 3, 733–736
37 Nygren, P-A. and Uhlén, M. (1997) Scaffolds for engineering novel binding sites in proteins. *Curr. Opin. Struct. Biol.* 7, 463–469
38 Hudson, P. (1998) Recombinant antibody fragments. *Curr. Opin. Biotechnol.* 9, 395–402

Interactions between prion protein isoforms: the kiss of death?

Byron Caughey

Direct interactions between the normal and aberrant forms of prion protein appear to be crucial in the transmission and pathogenesis of transmissible spongiform encephalopathies (TSEs) or prion diseases. Recent studies of such interactions *in vitro* have provided mechanistic insight into how TSE-associated prion protein might promote its own propagation in a manner that is specific enough to account, at least in part, for TSE strains and species barriers.

In its normal state, prion protein (PrP) seems innocent enough – a cell-surface glycoprotein that is present in most mammalian tissues. Occasionally, however, things go wrong with this protein and it accumulates in abnormal forms that cause devastating neurodegenerative diseases called transmissible spongiform encephalopathies (TSEs) or prion diseases[1,2]. TSE diseases include scrapie of sheep and goats, bovine spongiform encephalopathy (BSE or mad cow disease), human Creutzfeldt–Jakob disease (CJD), and chronic wasting disease (CWD) of deer and elk. A key feature of these diseases, besides the accumulation of abnormal PrP isoforms, is their transmissibility. Although the infectious agent, or prion, is not fully understood[3], substantial evidence suggests that it requires an abnormal, usually protease-resistant, form of PrP (Refs 2,4). Thus, Stanley Prusiner's proposal[2] that an abnormal form of PrP is the main component of the infectious agent remains a predominant, but controversial, working hypothesis in the TSE field.

J.S. Griffith originally proposed that a protein alone could serve as the infectious TSE agent if it were an abnormal, pathogenic form of a host protein that could induce its normal counterpart to convert to the abnormal form[5]. Since the discovery of PrP, evidence has mounted in favor of the importance of precise interactions between the normal and aberrant forms of PrP in TSE-agent propagation, transmission and pathogenesis[1,4,6]. Exciting recent advances in the understanding of TSE neuropathogenesis have been reviewed elsewhere[1]. This review will focus on the molecular interactions between the different PrP isoforms and how they might relate to TSE pathobiology.

Byron Caughey
Laboratory of Persistent Viral Diseases, NIAID, NIH, Rocky Mountain Laboratories, 903 S. 4th St, Hamilton, MT 59840, USA.
e-mail: bcaughey@nih.gov

PrP isoforms: normal versus abnormal

Normally, in its mature form after removal of N- and C-terminal signal sequences, cellular PrP (PrP^C) contains ~210 amino acid residues, two Asn-linked

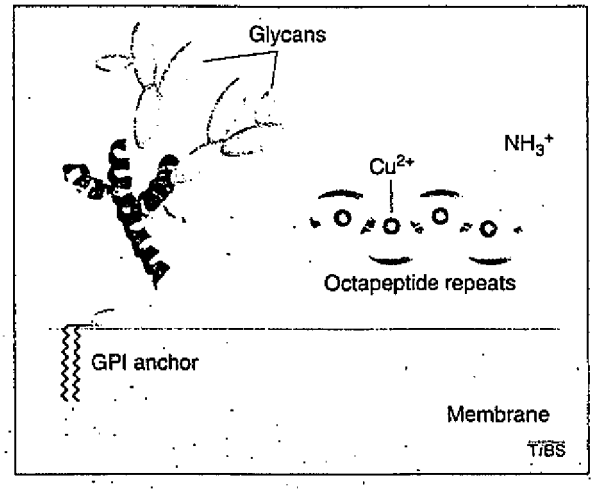


Fig. 1. Model of PrP^C structural domains. The folded C-terminal portion of PrP^C that contains the short β-sheet strands (yellow arrows) and the α helices (pink) is based on a model derived from the nuclear magnetic resonance (NMR)-based coordinates of residues 124–228 of hamster PrP (Ref. 36). The remainder of the molecule appears, by NMR, to be flexibly disordered. Abbreviations: GPI, glycophosphatidylinositol moiety; PrP^C, cellular prion protein.

http://tibs.trends.com 0968-0004/01/$ – see front matter. Published by Elsevier Science Ltd. PII: S0968-0004(01)01792-3

ELSEVIER

Journal of Immunological Methods 231 (1999) 25–38

JOURNAL OF IMMUNOLOGICAL METHODS

www.elsevier.nl/locate/jim

Single domain antibodies: comparison of camel VH and camelised human VH domains

Lutz Riechmann [a,*], Serge Muyldermans [b,1]

[a] MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

[b] Vlaams Interuniversitair Instituut Biotechnologie, Vrije Universiteit Brussel, Paardenstraat 65, B-1640 Sint Genesius Rode, Belgium

Abstract

The antigen binding sites of conventional antibodies are formed primarily by the hypervariable loops from both the heavy and the light chain variable domains. Functional antigen binding sites can however also be formed by heavy chain variable domains (VH) alone. In vivo, such binding sites have evolved in camels and camelids as part of antibodies, which consist only of two heavy chains and lack light chains. Analysis of the differences in amino acid sequence between the VHs of these camel heavy chain-only antibodies and VH domains from conventional human antibodies helped to design an altered human VH domain. This camelised VH proved, like the camel VH, to be a small, robust and efficient recognition unit formed by a single immunoglobulin (Ig) domain. Biochemical, structural and antigen binding characterisation properties of both camel VH domains and camelised human VH domains suggest that these can compete successfully with single chain variable domain (Fv) fragments from conventional antibodies in many applications. Of special importance in this respect is the use of such VH domains as enzyme inhibitors, for which they seem to be better suited than Fv fragments. This function appears to be closely related to their often very long third hypervariable loop, which is central for antigen recognition in their binding sites.

Keywords: Antibody engineering; Immunoglobulin; VH; Camel; Single domain; Heavy chain

1. Introduction

Antibodies have long been considered as a powerful tool to recognise and target almost any molecule with a high degree of specificity and affinity. For such purposes, natural antibodies can be obtained as mixtures in the form of antisera from immunised animals (Cohn et al., 1949) or as monoclonal antibodies from hybridomas (Köhler and Milstein, 1975). More recently, recombinant DNA technology has allowed the cloning and genetic manipulation of antibody genes, which can then be expressed as recombinant antibodies in eukaryotic cells (Boulianne

Abbreviations: CDR, complementarity determining region; CH, heavy chain constant domain; CL, light chain constant domain; ELISA, enzyme-linked immunoadsorbent assay; Fab, antigen binding fragment; FR, framework; Fv, variable domain fragment; g3p, gene 3 protein; H1 to H3, heavy chain hypervariable regions 1 to 3; Ig, immunoglobulin; NMR, nuclear magnetic resonance; PCR, polymerase chain reaction; scFv, single chain Fv; T_m, melting point; VH, heavy chain variable domain; VL, light chain variable domain

* Corresponding author. Tel.: +44-1223-402106; fax: +44-1223-402140; e-mail: lutz@mrc-lmb.cam.ac.uk

[1] E-mail: svmuylde@vub.ac.be.

0022-1759/99/$ - see front matter
PII: S0022-1759(99)00138-6

et al., 1984; Neuberger et al., 1985). To obtain good yields of active antibody protein also from bacteria, it was necessary to change the antibody format from the full-length molecule, consisting of two multidomain heavy and light chains each, to smaller versions like the antigen binding fragment (Fab), or variable domain (Fv) fragment (Better et al., 1988; Skerra and Plückthun, 1988). These fragments all contain ordinary antigen binding sites formed by a single pair of N-terminal variable domains from heavy and light chain, but they contain no (Fv) or only one constant domain (Fab) per chain. Apart from enabling better expression in bacteria, these smaller formats have other advantages as they, for example, result in a better distribution and faster clearance than larger antibody molecules when used in vivo (Yokota et al., 1990). Their smaller size makes them also more suitable for structural studies like nuclear magnetic resonance (NMR) spectroscopy (McManus and Riechmann, 1991; Riechmann et al., 1991).

It was therefore attempted to create even smaller fragments of antibodies with adequate antigen binding activities. To decrease size significantly below that of Fv fragments, such minimal recognition units have to be based on single domains. And, indeed, already very early experiments indicated that antibody heavy chains can occasionally bind antigens in the absence of their light chain partner (Utsumi and Karush, 1964). These results were corroborated, when single heavy chain variable (VH) domains were isolated from bacterial expression libraries of heavy chain variable regions from immunised mice (Ward et al., 1989).

2. Camel heavy chain antibodies

While these results suggested that it might well be possible to generate single domain antibodies in vitro, it was then discovered that heavy chain-only antibodies had also evolved in vivo. In camels and other camelid species, a significant proportion of the natural antibody repertoire were found to consist of antibodies lacking a light chain partner (Hamers-Casterman et al., 1993). These antibodies have a molecular weight of ~95 kDa instead of the ~160 kDa for conventional antibodies (Ungar-Waron et al., 1987). At least, two different fractions of heavy chain antibodies could be separated from the conventional heterotetrameric antibodies by differential adsorption on proteins A and G. Later on, the sequence of cDNA clones of spleen or blood lymphocytes revealed the presence of three or four different heavy chain antibody isotypes in dromedary and llama, respectively (Vu et al., 1997). These cDNA sequences lack the exon coding for the first constant domain. The heavy chain polypeptide of heavy chain antibodies is therefore composed of the variable domain, immediately followed by the hinge, CH2 and CH3 domains (Fig. 1). The absence of the CH1 domain explains the absence of the light chain in the heavy chain-only antibodies, as this domain is the anchoring place for the constant domain of the light chain (Padlan, 1994). Besides the presence of immunoglobulin (Ig) heavy chains without CH1, the analysis of the cDNA clones revealed that sera of camelids contain also two different γ-isotypes with a CH1 domain (Vu et al., 1997). Evidently, these latter encode the heavy chains of the conventional antibodies within the Camelidae.

A pathological disorder in humans or mice, known as heavy chain disease, occurring in humans or mice is characterised by the presence of heavy chain antibodies in their sera (Seligmann et al., 1979). These truncated antibodies result from a somatic event that removes various parts of the VH and CH1 region from the expressed Ig gene. The mouse or human heavy chain antibodies however are not func-

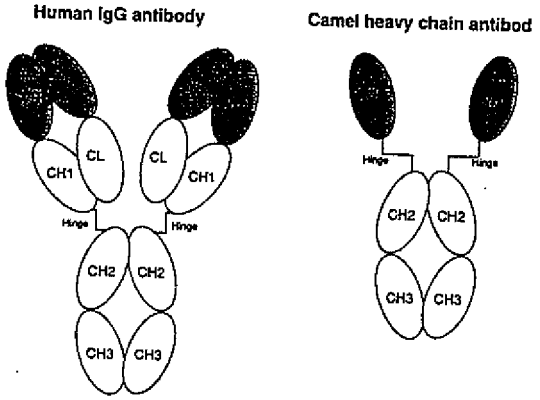


Fig. 1. Schematic picture of a conventional antibody (human γ-isotype) and a camel heavy chain antibody. The antigen binding site forming variable domains, which form either an Fv fragment in case of the conventional antibodies or a single domain antibody in case of the camel heavy chain antibodies, are shaded.

tional in antigen binding. In sharp contrast, the blood sampling of infected or immunised dromedaries showed that the immune response contained a diverse repertoire of heavy chain-only antibodies, which were functional in antigen binding (Hamers-Casterman et al., 1993).

3. Camelising human VH domains

Sequence analysis suggested that the camelid VH domains may contain a significantly altered surface in that region, which in VHs from conventional antibodies forms the VH/light chain variable domain (VL) interface (Chothia et al., 1985; Muyldermans et al., 1994). The antibody VH gene family, which is overall most homologous to VHs from camel heavy chain-only antibodies, is the human VH3 family (Muyldermans et al., 1994; Nguyen et al., 1998). Their framework (FR) sequences are very similar except for three residues in FR2 (Fig. 2), which are highly conserved in VH domains from most conventional antibodies. These residues (G44, L45, W47 in the human VH3; E44, R45, G47 in most camel VHs) are located in the VH/VL interface and the side chains of residues 45 and 47 point towards the VL in structures of conventional antibodies (Chothia et al., 1985).

To analyse the effect of these mutations, a human VH3 was expressed as an isolated domain in *Escherichia coli* and the three mutations were introduced in its former VL interface (Davies and Riechmann, 1994). The original human VH and two camelised mutants (G44E, L45R, W47I or W47G) were all found to be monomeric at low protein concentrations. However, the original human VH, in the absence of a VL domain, started to aggregate at higher protein concentrations (1 mg/ml and above) causing severe line broadening in NMR analysis (Davies and Riechmann, 1994). Aggregation was significantly reduced for the camelised VH domains. Thus, the camelised human VH domains had a much improved linewidth (transverse proton relaxation time T_2 was increased from 14.5 up to 29 ms) in NMR

```
FR1          1       5        10        15        20        25        30
Human VH3    E V Q L V E S G G G L V Q P G G S L R L S C A A S G F T F S
             . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Camel VH     E V Q L V E S G G G S V Q A G G S L R L S C A A S G Y T Y S

CDR1/FR2  31       35    36      40        45      49
Human VH3    @ @ @ @ @     W V R Q A P G K G L E W V S
                           . . . . . . . * * . * . .
Camel VH     @ @ @ @ G     W F R Q A P G K E R E G V S

CDR2      50  52 a53  55         60        65
Human VH3    @ @ @ @ @ @ @ @ @ @ Y A D S V K G
                               . . . . . . .
Camel VH     @ @ @ @ @ @ @ @ @ @ Y A D S V K G

FR3       66       70        75        80  82 a b c83  85        90       94
Human VH3    R F T I S R D N S K N T L Y L Q M N S L R A E D T A V Y Y C A R
             . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Camel VH     R F T I S Q D N A K N T V Y L Q M N S L K P E D T A I Y Y C A A

CDR3/FR4  95        100 102   103 105       110   113
Human VH3    @ @ @ @ @ @ @ @     W G Q G T L V T V S S
                                 . . . . . . . . . . .
Camel VH     @ @ @ @ @ @ @ @     W G Q G T Q V T V S S
```

Fig. 2. Consensus VH sequences from camel heavy chain-only antibodies (Muyldermans et al., 1994) and the human VH subgroup 3 (Kabat et al., 1991). The most frequent amino acid at each position is given except for residues 31 to 35, 50 to 58 and 95 to 102 (@), which are hypervariable in both VH families. Residues 44, 45 and 47 (*) were altered to camelise a human VH3 domain (see text). Residues identical in the camel VH and human VH sequence are indicated by a dot. FR1 to FR4 denote framework regions, CDR1, CDR2 and CDR3 denote complementarity determining regions and contain the hypervariable regions.

experiments, which enabled their structural analysis in solution (Davies and Riechmann, 1994; Riechmann, 1996; Riechmann and Davies, 1995). A very limited number of mutations in the former VL interface had therefore made it possible to create a human single VH domains, which behaved well in solution.

Apart from enabling the NMR analysis of the VH domain, this improvement opened the possibility to create single domain recognition units based on human heavy chain variable domains. Such camelised human VH domains may turn out to be more suitable than camel VH domains for any future in vivo use. They further carry, in case of the human VH3 domain, the additional advantage of being recognised by the bacterial superantigen protein A, which facilitates highly specific and efficient purification. Protein A binds to a nonlinear epitope within the VH (Riechmann and Davies, 1995) and thereby will result in purification of correctly folded VH protein only.

4. Preparation of specific camel heavy chain antibodies

The immunisation of llamas and dromedaries following standard protocols (complete and incomplete freund adjuvant; three to four boosts with 50 μg to 1 mg immunogen per animal) generates specific heavy chain-only antibodies of good titres. Approximately 0.1 mg of polyclonal heavy chain-only antibodies with specificity for α-amylase could be trapped by batch adsorption from 1 ml serum of an immunised dromedary (Lauwereys et al., 1998).

The polyclonal heavy chain-only antibodies of the IgG3-type purified on proteins A and G can also be used as a source to isolate variable domains. These variable domains are then obtained after a limited proteolytic digestion of IgG3 with endo-Glu V8 protease. This enzyme cleaves the short hinge region between the VH and CH2. Protein A chromatography retains the Fc containing fragments and a subfraction of the camel VHs. The flow-through contains the majority of the VH domains, of which measurable amounts could recognise the antigen (Lauwereys et al., 1998).

To obtain recombinant forms of camel VH domains, the repertoire of heavy chain-only antibody variable domains from an immunised camel are cloned in bacteria (Ghahroudi et al., 1997). Expression libraries of such VH repertoires can be screened for the presence of antigen specific binders. To avoid the contamination with VH genes originating from conventional antibodies a two-step polymerase chain reaction (PCR) was proposed (Ghahroudi et al., 1997). In the first PCR on cDNA template of blood lymphocytes a set of primers, which anneals at the first codons of the VH and within the CH2 region, were used to amplify all γ-isotypes. The PCR products originating from the heavy chain-only antibodies could be selectively eluted after gel electrophoresis since the fragment of the heavy chain-only antibody γ-isotypes are ~350 nucleotides shorter due to the absence of the CH1 exon. The entire VH region is then reamplified with nested primers annealing at the codons of FR1 and FR4 of the VH, respectively. The resulting PCR product is finally ligated in a phage display vector (see below) adapted to clone a VH region only. A VH library of 10^6–10^7 individual clones, from which typically several antigen binders were selected, can be obtained from as little as 5–10 ml blood. The whole procedure from the first immunisation to the identification of binders can be performed in less than 3 months. Following this procedure, more than 20 binders against a variety of proteinaceous antigens were isolated (Table 1). These VHs bound their antigen specifically with dissociation constants from 100 nM into the subnanomolar range (Muyldermans and Lauwereys, 1999). This affinity is similar to that measured for conventional antibodies from a secondary immune response (Foote and Milstein, 1991).

Although phage display or expression libraries for Fv or Fab fragments of conventional antibodies can also be prepared from blood of immunised or infected mice and humans (Persson et al., 1991; Winter et al., 1994), the single domain VH libraries of immunised camels have several advantages. First, all the variable domains of camelid heavy chain antibodies belong to one single family (VH3). Such a VH library can therefore be generated with a single set of PCR primers. Secondly, the entire paratope is located on a single domain present in a single exon that is cloned as one entity. In contrast, the VH genes of mouse and human antibodies belong to several families so that multiple PCRs with different sets of primers are necessary to clone the complete

Table 1
Affinities of camel VH domains and camelised human VH domains
Data for the camel VH are from Ghahroudi et al. (1997), Lauwereys et al. (1998) and K. Conrath (personal communication), data for the camelised human VHs are from Davies and Riechmann (1996b). The antigen used were α-amylase, carbonic anhydrase, β-lactamase, 4-glycyl-2-phenyloxazol-5-one (OxGly), 3-iodo-4-hydroxy-5-nitrophenyl-acetyl caproic acid (NIP-CAP), the HIV transcription factor rev and hen egg lysozyme. The length of the H3 region refers to the number of residues between 95 and 102 inclusive using the Kabat nomenclature (Kabat et al., 1991).

	Ligand	K_d (nM)	H3-length
Camel VH			
cAb-AMD7	α-amylase	15	16
cAb-AMD9	α-amylase	3.5	14
cAb-AMB10	α-amylase	24	14
cAB-CA04	carbonic anhydrase	29	18
cAb-CA06	carbonic anhydrase	20	13
cAb-Lys2	lysozyme	2	19
cAb-Lys3	lysozyme	65	24
cAb-βla01	β-lactamase	< 1	21
cAb-βla02	β-lactamase	1.5	8
cAb-βla03	β-lactamase	10	17
Camelised VH			
VH-Ox21	OxGly	146	15
VH-OS1	OxGly	25	15
VH-Ox21.2.4	OxGly	47	15
VH-Ox62	OxGly	267	15
VH-N3c1	NIP-CAP	292	15
VH-N3c1.2.2	NIP-CAP	31	15
VH-REVg1	rev	220	11
VH-REVrs1	rev	401	11
VH-LS2	lysozyme	3100	10
VH-LS2.5.1	lysozyme	1600	10

repertoire. In addition, the Fv has to be reconstituted from the combination of a VH and a VL domain. These must be amplified as separate gene fragments, which are randomly combined afterwards. Therefore, original pairs of VH and VL domains with antigen binding activity that were matured during the immune response as one entity can only be recovered by random combination requiring Fv libraries of relatively large size, which at times approaches the limitations of bacterial transformation efficiency.

5. Preparation of specific camelised VH domains

While camel heavy chain-only antibodies of a desired antigen specificity can be rescued from natural sources, camelised VH domains are synthetic and must therefore be generated in vitro. For this, the human VH3 domain with the camelised VL interface (i.e., with mutated residues 44, 45 and 47) was used as a building block for the creation of designed libraries. VH repertoires can most simply be created through the introduction of randomised regions into the VH gene to vary the hypervariable loops both in length and amino acid residue nature. The three hypervariable loops of VH and VL form the antigen binding site in ordinary antibodies (Chothia et al., 1989).

To facilitate selection of a repertoire of camelised VH domains, these were displayed, like other antibody fragments before, on filamentous phage by fusion to the N-terminus of the minor phage coat gene 3 protein (g3p) on the gene level (Davies and Riechmann, 1995a). Each phage particle displays its individual VH, which is encoded in the encapsulated phage genome. VHs can therefore be selected, for several rounds if needed, through panning of phage on immobilised antigen. Bound phage is rescued and regrown through infection of bacteria. Selected VHs can be analysed (and indeed be used for detection or targeting) for their antigen specificity and affinity as phage displayed VH or as soluble VH. Soluble VH can be produced after either subcloning into a soluble expression vector or through the use of bacterial suppresser strains and appropriate stop codons between the VH fusion and the phage g3p (Hoogenboom et al., 1991).

Repertoires of camelised VH domains were initially created by randomisation of residues within the third hypervariable loop H3, which at the same time was varied in length. In VH domains from conventional antibodies and indeed camel heavy chain-only antibodies, this loop contains among the three VH hypervariable loops the highest diversity in length and amino acid nature (Kabat et al., 1991; Wu et al., 1993). From such a repertoire of 2×10^8 clones, camelised VH domains specific for hapten, peptide and protein antigens were selected (Davies and Riechmann, 1995a, 1996a; Martin et al., 1997). The dissociation constants for these VH domains in their soluble form and their respective antigens were in the nanomolar to micromolar range (Table 1). The affinities could be improved up to tenfold by phage selection of VHs containing secondary mutations

within the other two hypervariable loops (Davies and Riechmann, 1996b).

These results show that camelised VH domains present a useful building block for the preparation of Ig-based recognition units of minimal size. Selected VH domains proved to be highly specific and of reasonable affinity, which can be increased through the preparation of larger repertoires or secondary mutations.

The same experiments, however, also made obvious that some properties of camelised VH domains can be improved upon. These are related to the former VL interface of the VH. The three camelising mutations (G44E, L45R and W47G), which increased the solubility of the single human VH domain most and lead to the least nonspecific binding of phage displayed VH domains, compromised the stability and most significantly the expression yield of active protein for the resulting soluble VH (Davies and Riechmann, 1994, 1995a). Camelised VH domains, which had been selected with a glycine at position 47, were therefore expressed as G47I mutants in their soluble version. This had in most cases minor or no effects on antigen affinity. The G47I mutation increased in all cases stability and purification yields of active VH protein significantly. Unfortunately, constitutive use of isoleucine at position 47 of the VH for phage display lead to an unacceptably high number of nonspecifically binding clones in the repertoire, which compromised selection (Davies and Riechmann, 1995a).

Further mutations in the camelised, human VH domain appear to be necessary to create a prototype FR, which exhibits low nonspecific binding and high expression yields at the same time. Thus, it was indeed possible to improve protein stability of camelised VH domains significantly through the introduction of a new intradomain disulphide bridge between cysteines at positions 33 (in FR2) and 100b (in H3), which could even be constitutively introduced in repertoires of camelised VH domains (Davies and Riechmann, 1996a).

Another problem concerns the multimerisation state of the selected VH domains. While most selected camelised VH domains were monomeric, dimeric domains also were occasionally selected (Martin et al., 1997). This obviously depends on the nature of the H3 loop as this is the only portion of the camelised VHs, which is different among the VH domains in the original repertoire.

Concerning the generation of diversity within repertoires, increased variation both in length and sequence of the H3 loop in camelised VH domains is probably the most effective and simplest method. Sequences from antigen specific, natural camel heavy chain antibodies after immunisation suggest that the H3 loop has an even more central role for antigen binding in heavy chain-only antibodies than it has in conventional antibodies (Muyldermans et al., 1994). Repertoires based on completely randomised H3 loops can be further diversified through a limited number of mutations in the other two hypervariable loops.

6. Expression of recombinant heavy chain-only antibodies and single VH domains

6.1. Phage antibody fragments

For the phage display of antibodies, these are typically fused as Fab or single chain Fv (scFv) fragments to the phage g3p protein. Once a specific antibody clone has been selected, this can be used for many applications directly in this phage format (Nissim et al., 1994). For example, selected phages can be most easily screened from supernatants of infected bacteria in enzyme-linked immunoadsorbent assay (ELISA) for antigen binding, where they are detected with an anti-phage antibody (e.g., anti-M13 monoclonal antibody; Pharmacia). This essay is highly sensitive as the phage coat contains about 2700 copies of the major phage coat protein gene 8 protein (Model and Russel, 1988), which forms the epitope for the anti-phage antibody. When a pure phage system (rather than a phagemid in combination with a helper phage) is used to provide the g3p-fusion, the phage contains up to five copies of the antibody fragment resulting in avidity effects, which will allow the detection of even very weak binders.

Concerning the use of phage displayed antibody fragment, there is no difference between ordinary heavy and light chain antibody fragments and single chain polypeptide antibodies like the camel heavy chain-only antibodies or camelised human VH do-

mains. VHs from camel heavy chain antibodies and the camelised human VH domains have both been successfully fused to g3p (Davies and Riechmann, 1995a; Ghahroudi et al., 1997) and were suitable for detection of their antigens in ELISA. Their use in other application, like histological staining or cell sorting, has not been tried, but there is no reason why they should perform any worse or better than ordinary antibodies.

6.2. *Soluble antibody fragments of camel heavy chain-only antibodies*

The genes for camel VH domains can be inserted by cassette mutagenesis into suitable expression vectors to produce larger multidomain or multifunctional proteins. For example, an intact heavy chain-only antibody was readily generated by cloning a particular camel VH in front of the hinge and effector function domains of human IgG1. This construct can be expressed in bacteria, but expression in mammalian cell lines will add the proper glycosylation at the CH2 domain. Such constructs in a pcDNA3 vector (Invitrogen) produced after transient expression in COS cells some 5 mg/l medium (K.B. Vu, personal communication). These chimeric heavy chain antibodies were fully active in antigen binding.

Multimerised forms of camel VHs linked on a single gene through a peptide were also obtained from bacteria yielding multivalent recognition units (K. Conrath, personal communication). Two camel VHs directed against different antigens should form bispecific recognition units when linked on the gene level. Equivalent constructs have been made using conventional scFv as molecular building blocks (Neri et al., 1995). However, the single domain nature of the camel VH will most likely yield higher functional expression levels.

We are convinced that camel VH will be most useful in diagnostic applications where stable and, small reporter molecules are required. Camel VHs can routinely be expressed with a His tail in the bacterial periplasm. The average purification yield obtained after purification by immobilised metal affinity chromatography and gel filtration is between 0.5 and 5 mg/l culture. These recombinant camel VHs do not dimerise and can be concentrated to 10 mg/ml without aggregation. The VHs are stable, they resist incubations at 37°C for 1 week and T_m values between 60°C and 72°C were measured (Ghahroudi et al., 1997).

6.3. *Soluble camelised human VH domains*

Camelised VH domains, which were selected from phage display libraries, have been expressed as soluble VH domains after subcloning into *E. coli* expression vectors directing folded protein to the periplasm through fusion to a suitable leader peptide (Davies and Riechmann, 1995a). Camelised VHs containing a glycine at position 47, which was most successfully used for selection after phage display, yielded up to 1 mg functional protein from 1-l bacterial culture grown in shaker flasks using protein A sepharose for purification from both supernatants or periplasmic preparations. The same VH domains, expressed with an isoleucine at position 47, yielded up to 10 mg active protein from 1 l of culture. The same was usually the case even when the VHs contained two additional cysteines (Davies and Riechmann, 1996a). ScFvs rescued from phage display libraries or indeed engineered from monoclonal antibodies have often similar purification yields (Skerra, 1993), which however can be highly variable depending on the particular scFv used. This is due to the wider structural variety of scFvs, while most camelised VHs, which are identical except for their H3 loop, give purification yields in the discussed range. In addition, expression of both camel VH domains and camelised human VH domains is not compromised by the presence of a linker between VH and VL as in a scFv. Such linkers can interfere with folding and can be susceptible to proteolysis (Whitlow et al., 1993).

Concerning the stability of the camelised VH domains, these are usually considerably higher than those of scFv or ordinary Fv fragments, which can fall more easily apart due to their two-domain architecture. The camelised VH domains containing a glycine at position 47 have melting points of about 60°C, while those with an isoleucine at position 47 have a T_m of about 70°C (Davies and Riechmann, 1995b, 1996a). Other mutations, like the change of valine 37 to phenylalanine and the introduction a second intradomain disulphide bond between newly

introduced cysteines at positions 33 and 100b, which were also adapted from natural camel VH domains, increased the thermostability of the VH domains to almost 80°C (Davies and Riechmann, 1996a). Again, as with purification yields, stabilities of scFvs are much more variable due to their more diverse nature of underlying VH and VL genes in addition to the variable lability of the VH/VL interaction.

Other features of camelised VH domains are less straightforward to compare with those of other formats of antibody fragments. Thus, camelised VH domains with long third hypervariable loops were occasionally cleaved by bacterial proteases. Although this may also be occasionally observed in scFvs, it will be less frequent in the case of conventional antibody combining sites as these contain usually shorter H3 loops.

Another observation is the formation of dimeric, camelised VH domains (Martin et al., 1997). Dimerisation is also not unknown for scFvs. In the case of scFvs, it was possible to turn that into a feature rather than a problem. Through the use of short VH/VL linkers, dimerisation (and even trimerisation) can even be enforced and can lead to so-called diabodies with two identical binding sites, or even different specificities when VH and VL domains from two different scFvs are linked on the same gene (Holliger et al., 1993). In the case of single VH domains, dimerisation through covalent linkage of two different VHs should also lead to avidity effects or to the creation of dual specificity. However, such modifications will compromise the two most important advantages of VH domains, which are small size and high stability.

As far as aggregation of camelised VH domains compared to other antibody fragments is concerned, this will very much depend on the nature of the particular fragment. Most camelised VH domains or Fv and Fab fragments do not aggregate, however, occasionally fragments are prone to aggregation depending on particular sequences within the proteins.

7. Structures of camel VH domains

The crystal structures of one llama and two dromedary VH domains are available in the protein data bank. The llama VH with specificity for human chorionic gonadotropin was solved to 1.8 Å resolution (pdb-file 1HCV) (Spinelli et al., 1996). The dromedary VHs, cAb-Lys3 (pdb-file 1MEL) cAb-RN05 (pdb-file 1BZQ) and cAb-RN05 (pdb-file 1BZQ), were crystallised with their respective antigens, hen egg-white lysozyme and bovine RNase A (Desmyter et al., 1996; Decanniere et al., 1999; Fig. 3). In all three cases, the Ig fold of the VH is well preserved. Two β-sheets (one with four and one with five β-strands) are packed against each other and stabilised by a conserved intradomain disulphide bond between C22 and C92. The rms deviation between the core of any two camel or llama VH and human VH (of family 3) is between 0.27 and 0.69 Å (Decanniere et al., 1999).

The side of the camel VH domain corresponding to the VL interacting face of the normal VH in an Fv has a quite different architecture. Compared to the human VH, four amino acid substitutions (V37F, G44E, L45R and W47G or W47S) are located in this region. The substitutions at positions 44, 45 and 47 were those, which were used to camelise the human VH and rendered the isolated domain more soluble (Davies and Riechmann, 1994). The nonpolar to polar amino acid substitutions (G44E and L45R) increase the hydrophilicity of the surface. The substitutions at positions 37 and 47 cause a net shift of the bulky hydrophobic groups. In case of the two camel VHs with a known structure, the H3 loop folds over these residues and makes them solvent inaccessible (Decanniere et al., 1999).

From a survey of all human and mouse VH antigen binding loop structures, it became apparent that only a restricted number of possible conformations were encountered (Chothia and Lesk, 1987; Chothia et al., 1992; Al-Lazikani et al., 1997). Three and four different conformations are described for the first and second antigen binding loop, respectively. These so-called canonical structures are determined by the length of the loop and the presence of particular residues at key positions. The H3 loop is extremely variable in length and sequence (Wu et al., 1993). Consequently, the prediction of the loop architecture remains more speculative despite recent progress (Martin and Thornton, 1996; Shirai et al., 1996; Morea et al., 1998). Surprisingly, the antigen binding loop structures of camel VHs deviate from the present canonical loop definitions of human and

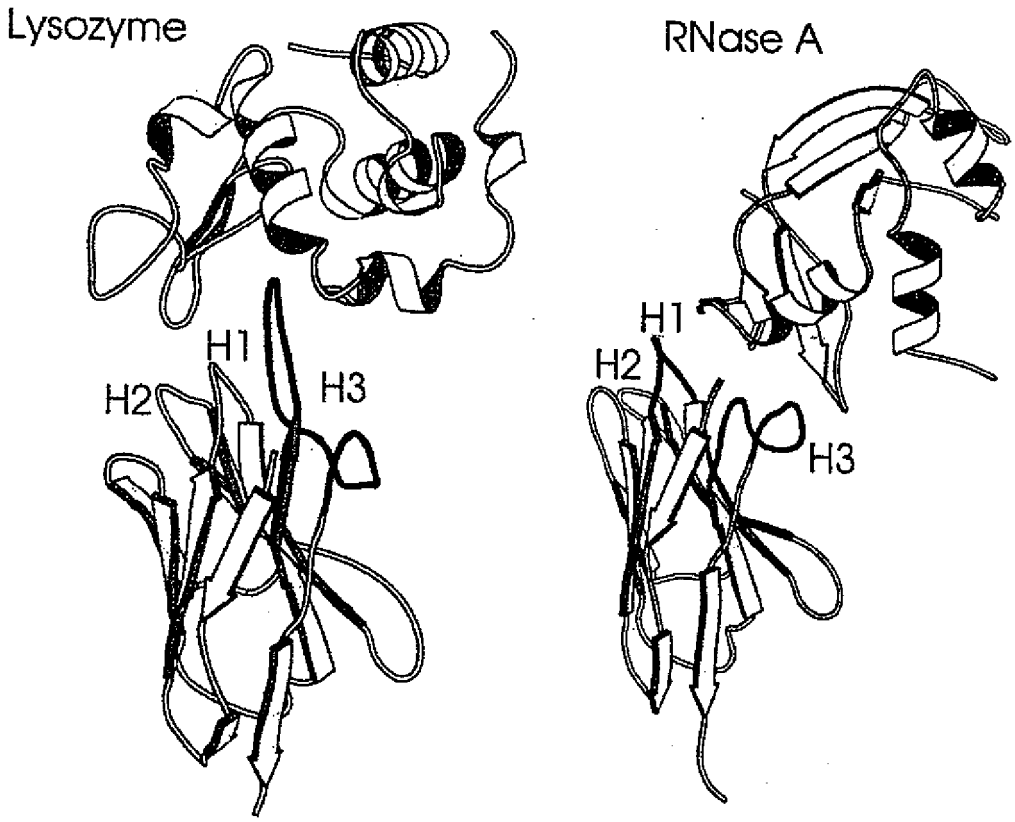


Fig. 3. Crystal structures of the complexes between two camel VH domains and their antigens. The complex cAb-Lys3::lysozyme (left) illustrates how the long H3 loop (black) protrudes from the remaining paratope and inserts into the catalytic site of lysozyme. In the cAb-RN05::RNase A model (right) only two VH loops, H1 and H3 (black) are involved in antigen binding. The H3 loop of the cAb-RN05 antibody (12 amino acids) is only half of that of cAb-Lys3. The H3 loop of cAb-RN05 does not protrude from the remaining paratope but folds over the former VL interface.

mouse VHs. This deviation could not be predicted, because the loop length and the residues at the key positions are very similar between camel VH and human VH. Consequently, the prediction of the camel VH hypervariable loop structures by the current canonical loop structure algorithms is not reliable. The additional canonical loop structures in camel VHs make the structural repertoire of their paratope larger than that of VH domains in Fv fragments from conventional antibodies.

Moreover, the hypervariable region around the first antigen binding loop in VHs of camel heavy chain-only antibodies is enlarged to cover residues 26 to 35 (Vu et al., 1997). In human VHs, the first hypervariable region only comprises positions 31 to 35 (Kabat et al., 1991). This indicates that residues 26 to 30 in camel VHs may more frequently participate in antigen binding. Indeed residue 29 (I29) of cAb-Lys3 interacts directly with the antigen lysozyme (Desmyter et al., 1996) and the hydroxyl group of Y27 of cAb-RN05 is hydrogen bonded with its RNase A antigen (Decanniere et al., 1999). The first amino acid of conventional VHs, which usually interacts with antigen, is located at position 30 (Tomlinson et al., 1996). The extension of the first hypervariable region in camel VHs and a concomitant enlarged antigen binding surface compared to that of a VH in a conventional antibody appears to compensate in part for the absence of a VL domain.

However, the antigen binding surface of cAb-RN05 remains small (570 Å^2) compared to Fv paratopes interacting with proteins (613–841 Å^2) due partly to a lack of participation of the H2 loop in antigen binding (Fig. 3). A large number of main

chain hydrogen bonds and van der Waals contacts leads nevertheless to an dissociation constant with antigen of 30 nM.

The cAb-Lys3 has a very long H3 loop of 24 residues, of which the first 10 protrude from the antigen binding site (Fig. 3). As a consequence, the actual antigen binding surface (847 Å^2) becomes even larger than that of most conventional Fvs (613–841 Å^2). The protruding H3 loop is constraint by a disulphide bond towards the H1 loop and by the parallel stacking of two tyrosines in the interior of loop H3 (Desmyter et al., 1996).

No antigen binding data are available for the llama VH. However, its short H3 loop (Spinelli et al., 1996) suggests that the actual antigen binding surface will be correspondingly small.

8. Structure of a camelised VH domain

The structure of one camelised human VH domain was solved in solution by NMR spectroscopy (pdb-file 1VHP). This VH contains three camelising mutations (G44E, L45R and W47I) in its former VL interface. The overall β-sheet structure is very similar to that of a noncamelised, VL-associated human VH3 domain (Pot; Fan et al., 1992) with a C^{α} rms of 2 Å for 82 (out of 113) aligned residues (Riechmann, 1996). The first two hypervariable loops of the camelised VH and the Pot-VH have the same length and adopt the same canonical structures. The H3 loop folds in the Pot-VH towards the VL and participates in VH/VL interactions. The (by four residues shorter) H3 loop in the camelised VH in contrast is orientated more towards the H1 and H2 loops. This is also in conflict with the camelid VH domains where the H3 loop, or part of the H3 loop, covers the side of the VH, which would form the VL interface in conventional antibodies. In the camelid VHs, the H3 loop covers hydrophobic regions of the VH, which would otherwise be exposed to the solvent. The camelising mutations reduce the hydrophobic character of the protein surface both in the camel and the camelised VHs, as three hydrophobic residues (44, 45 and 47) are replaced by more hydrophilic amino acids. Hydrophobicity of the former VL interface in the camelised VH was further reduced by the reorientation (compared to a VL-associated VH) of the side chains of the nonmutated residues 37, 38 and 103. The hydrophobic side chains of residues V37 and W103 are completely or partly buried in the camelised VH, while they point towards to the VL in a VL-associated VH. The opposite is the case for the hydrophilic R38 side chain. These changes were not found in the camel VH structures and indicate a principle structural difference between camel and camelised VH domains. Whether this will also be true for camelised VHs with extremely long H3 loops as in natural camel VHs remains to be seen. Folding back of a long H3 loop onto the former VL interface in camelised VH domains might well influence its structure.

No direct structural details are available for antigen binding by camelised VH domains, but affinity improvement due to changes in the H1 and H2 loops suggest that these may also be involved in antigen binding (Davies and Riechmann, 1996b). The H3 loop will almost certainly play a central role in their binding site as it formed the source of structural variety in the underlying repertoire, from which the VH domains were selected.

9. Special features of single domain antibodies

The observation of a convex paratope architecture in the camel VH cAb-Lys3 (Desmyter et al., 1996) was most remarkable since it has never been observed in Fvs (Padlan, 1996). Normally, the antigen binding loops of an Fv form a cavity, a groove or a flat surface (Webster et al., 1994). These Fv topographies are correlated with the size and type of antigen. Haptens tend to be bound into cavities of the paratope, peptides bind into a groove and larger antigens such as proteins are bound by antibodies with flat paratopes, eventually undulated by some side chains to improve the complementarity of the surfaces (Padlan, 1996). Large protruding loops of 10 amino acids or more seem to be unique for the antigen binding sites of heavy chain-only antibodies. This feature gives camel VHs a special niche for antigen recognition, as their long third hypervariable loop can insert into cavities of antigen surfaces. This is especially important, as the catalytic site of en-

zymes is often located at the largest cavity on their protein surface (Laskowski et al., 1996). Such sites are usually not immunogenic for conventional antibodies (Novotny et al., 1986). Camel heavy chain-only antibodies however are able to recognise such epitopes.

Indeed, in the structure of camel VH cAb-Lys3, the 24 residue long H3 loop penetrates deeply into the active site of lysozyme (Transue et al., 1998) suggesting that camel heavy chain antibodies might be able to form specific enzyme inhibitors. Proof was given by the successful retrieval of specific inhibitors from the VH library of a camel immunised with α-amylase, lysozyme and carbonic anhydrase (Lauwereys et al., 1998). Similarly, from a library of camelised VHs with randomised third hypervariable loops (Davies and Riechmann, 1995a) an inhibitor for the hepatitis C virus NS3 protease could be selected in vitro (Martin et al., 1997).

Whether inhibition is indeed in all cases due to the insertion of long hypervariable loops into active site cavities of the enzymes will have to be confirmed by structural studies, and it cannot be excluded that inhibitory VHs might have additional strategies to inactivate enzymes.

The paratope of camel VHs comprises only three antigen binding loops, of which H3 provides most of the contacts. The lower complexity of their antigen binding site might therefore make single domain antibodies a suitable FR for the peptide scanning technique (Laune et al., 1997) to design smaller peptides or peptide analogues with enzyme inhibiting or receptor blocking capacity derived from the sequences of the H3 loops.

10. Conclusions

The most basic difference between camel VHs and camelised VHs is that camel VHs have naturally evolved and can be obtained by in vivo immunisation. While immunisation can be mimicked in vitro through selection of synthetic phage displayed VH domains, the architecture of the camel VHs is at present still superior to that of camelised VHs because of a more mature design as a result of their natural evolution. Camelised human VH domains may need additional modifications in their now exposed, former VL interface to tolerate the exposure to a hydrophilic surrounding while maintaining a high stability and good folding properties and providing a highly diverse antigen binding site. Once this disadvantage is overcome, camelised and camel VH domain should be equally well suited for the in vitro selection of antigen specific single domain antibodies. Concerning the in vivo selection of VH domains with good binding properties after immunisation of animals, camel heavy chain-only antibodies will obviously remain unchallenged. Advantages of camelised, human VHs compared to camel VH domains however include the presence of a Protein A binding in the case of domains based on the human VH3 gene family and the possibly more attractive use of camelised, human VHs for any therapeutic purpose in humans due to a probably lower immunogenicity.

Given a camel available for immunisation and an antigen ready for injection, it is relatively straightforward to generate a recombinant camel single domain antibody with good affinity and specificity. This may prove particularly valuable when specific enzyme inhibitors are needed, as these seem to be very rare among conventional antibodies but frequent among camel heavy chain antibodies. This property appears to be closely related to their often very long H3 loop. This niche for antigen binding together with the advantages in size and stability should lead to an important role for single VH domains for biotechnological applications in the future.

It seems therefore also desirable to establish in vitro systems for the generation of such single domain antibodies avoiding the need for animals. To prepare single domain antibodies based on natural camel VHs in vitro, a large bank of the VH germline segments, which are used in the camel heavy chain-only antibodies, must be prepared and for example completed with synthetic H3 loops and a suitable C-terminal FR. Their display on phage will allow in vitro selection of specific recombinant VH domains. Camelised, human VH domains present a viable alternative to camel VHs. These however may still require some improvements to efficiently compete with the more reliable biochemical properties of natural camel VH domains. Such a design may eventually lead to domains even smaller than full-

length VH domains, as already shown in the case of the VH-based β-domains (Pessi et al., 1993; Martin et al., 1994). Alternatively, it may also be possible to utilise light chain variable domains for this purpose, although these lack at least in nature the feature of a frequently very long hypervariable loop.

In any case, single VH domains have so far proven to perform no worse than Fv or Fab fragments from ordinary antibodies concerning antigen binding and purification yields. In addition, single VH domains are definitely more stable due to their simpler architecture. They are therefore likely to challenge conventional antibodies for many biotechnological applications, while in most therapeutic applications conventional antibodies will remain the reagent of choice for reasons of immunogenicity. Possible exceptions even in therapy are antibody-based enzyme inhibitors, as only camel heavy chain-only antibodies seem to perform this function in vivo, and anti-idiotypic vaccinations, as camel VHs should raise good responses in noncamelid species due to their noncommon structures of antigen binding loops.

References

Al-Lazikani, B., Lesk, A.M., Chothia, C., 1997. Standard conformations for the canonical structures of immunoglobulins. J. Mol. Biol. 273, 927.

Better, M., Chang, C.P., Robinson, R.R., Horwitz, A.H., 1988. *Escherichia coli* secretion of an active chimeric antibody fragment. Science 240, 1041.

Boulianne, G., Hozumi, N., Shulman, M.J., 1984. Production of functional chimeric mouse/human antibody. Nature 312, 643.

Chothia, C., Lesk, A.M., 1987. Canonical structures for the hypervariable regions of immunoglobulins. J. Mol. Biol. 196, 901.

Chothia, C., Novotny, J., Bruccoleri, R., Karplus, M., 1985. Domain association in immunoglobulin molecules. The packing of variable domains. J. Mol. Biol. 186, 651.

Chothia, C., Lesk, A.M., Tramontano, A., Levitt, M., Smith-Gill, S.J., Air, G., Sheriff, S., Padlan, E.A., Davies, A.H., Tulip, W., Colman, P.M., Spinelli, S., Alzari, P.M., Poljak, R., 1989. Conformations of immunoglobulin variable regions. Nature 342, 877.

Chothia, C., Lesk, A.M., Gherardi, E., Tomlinson, I.M., Walter, G., Marks, J.D., Llewelyn, M.B., Winter, G., 1992. Structural repertoire of human VH segments. J. Mol. Biol. 227, 799.

Cohn, E.J., Strong, L.E., Hughes, W.L., Mulford, D.J., Ashworth, J.N., Melin, M., Taylor, H.L., 1949. Preparation and properties of serum and plasma proteins: IV. J. Am. Chem. Soc. 68, 459.

Davies, J., Riechmann, L., 1994. Camelising human antibody fragments: NMR studies on VH domains. FEBS Lett. 339, 285.

Davies, J., Riechmann, L., 1995a. Antibody VH domains as small recognition units. Bio/Technology 13, 475.

Davies, J., Riechmann, L., 1995b. An antibody VH domain with a lox-Cre site integrated into its coding region: bacterial recombination within a single polypeptide chain. FEBS Lett. 377, 92.

Davies, J., Riechmann, L., 1996a. Single antibody domains as small recognition units: design and in vitro antigen selection of camelised, human VH domains with improved stability. Protein Eng. 9, 531.

Davies, J., Riechmann, L., 1996b. Affinity improvement of single antibody VH domains. Residues in all three hypervariable regions affect antigen binding. Immunotechnology 2, 169.

Decanniere, K., Desmyter, A., Lauwereys, M., Ghahroudi, M.A., Muyldermans, S., Wyns, L., 1999. A single-domain antibody fragment in complex with RNase A: noncanonical loop structures and nanomolar affinity using two CDR loops. Struct. Folding Design 4, 361.

Desmyter, A., Transue, T.R., Arbabi Ghahroudi, M., Dao-Thi, M.-H., Poortmans, F., Hamers, R., Muyldermans, S., Wyns, L., 1996. Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme. Nat. Struct. Biol. 3, 803.

Foote, J., Milstein, C., 1991. Kinetic maturation of an immune response. Nature 352, 530–532.

Ghahroudi, M.A., Desmyter, A., Wyns, L., Hamers, R., Muyldermans, S., 1997. Selection and identification of single domain antibody fragments from camel heavy-chain antibodies. FEBS Lett. 414, 521.

Hamers-Casterman, C., Atarhouch, T., Muyldermans, S., Robinson, G., Hamers, C., Bajyana Songa, E., Bendahman, N., Hamers, R., 1993. Naturally occurring antibodies devoid of light chains. Nature 363, 446.

Holliger, P., Prospero, T., Winter, G., 1993. Diabodies: small bivalent and bispecific antibody fragments. Proc. Natl. Acad. Sci. U.S.A. 90, 6444.

Hoogenboom, H.R., Griffiths, A.D., Johnson, K.S., Chiswell, D.J., Hudson, P., Winter, G., 1991. Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. Nucleic Acids Res. 19, 4133.

Kabat, E., Wu, T.T., Perry, H.M., Gottesman, K.S., Foeller, C., 1991. Sequence of proteins of immunological interest. US Public Health Services, NIH Bethesda, MD, Publication No. 91.

Köhler, G., Milstein, C., 1975. Continuous cultures of fused cells secreting antibody of predefined specificity. Nature 256, 495.

Laskowski, R.A., Luscombe, N.M., Swindells, M.B., Thornton, J.M., 1996. Protein clefts in molecular recognition and function. Protein Sci. 5, 2438.

Laune, D., Molina, F., Ferrieres, G., Mani, J.-C., Cohen, P., Simon, D., Bernardi, T., Piechaczyk, M., Pau, B., Granier, C., 1997. Systematic exploration of the antigen binding activity of synthetic peptides isolated from the variable regions of immunoglobulins. J. Biol. Chem. 272, 30937.

Lauwereys, M., Ghahroudi, M.A., Desmyter, A., Kinne, J., Hölzer, W., De Genst, E., Wyns, L., Muyldermans, S., 1998. Potent enzyme inhibitors derived from dromedary heavy-chain antibodies. EMBO J. 17, 3512.

Martin, A.C.R., Thornton, J.M., 1996. Structural families in loops of homologous proteins: automatic classification, modelling and application to antibodies. J. Mol. Biol. 263, 800.

Martin, F., Toniatti, C., Salvati, A.L., Venturini, S., Ciliberto, G., Cortese, R., De Fransesco, R., Sollazzo, M., 1994. The affinity selection of a minibody polypeptide inhibitor of human interleukin-6. EMBO J. 13, 5303.

Martin, F., Volpari, C., Steinkühler, C., Dimasi, N., Brunetti, M., Biasiol, G.A.S., Cortese, R., De Fransesco, R., Sollazzo, M., 1997. Affinity selection of a camelised VH domain antibody inhibitor of hepatitis C virus NS3 protease. Protein Eng. 10, 607.

McManus, S., Riechmann, L., 1991. Use of 2D NMR, protein engineering and molecular modeling to study the hapten-binding site of an antibody Fv fragment against 2-phenyloxazolone. Biochemistry 30, 5851.

Model, P., Russel, M., 1988. Filamentous bacteriophage. In: Calender, R. (Ed.), The Bacteriophages. Plenum, New York, p. 375.

Morea, V., Tramontano, A., Rustici, M., Chothia, C., Lesk, A.M., 1998. Conformations of the third hypervariable region in the VH domain of immunoglobulins. J. Mol. Biol. 275, 269.

Muyldermans, S., Lauwereys, M., 1999. Unique single-domain antigen binding fragments derived from naturally occurring camel heavy-chain antibodies. J. Mol. Recognit. 12, 131.

Muyldermans, S., Atarhouch, T., Saldanha, J., Barbosa, J.A.R.G., Hamers, R., 1994. Sequence and structure of VH domain from naturally occurring camel heavy chain immunoglobulins lacking light chains. Protein Eng. 7, 1129.

Neri, D., Momo, M., Prospero, T., Winter, G., 1995. High affinity antigen binding by chelating recombinant antibodies (CRAbs). J. Mol. Biol. 246, 367.

Neuberger, M.S., Williams, G.T., Mitchell, E.B., Jouhal, S.S., Flanagan, J.G., Rabbitts, T.H., 1985. A hapten-specific chineric IgE antibody with human physiological effector function. Nature 314, 268.

Nguyen, V.K., Muyldermans, S., Hamers, R., 1998. The specific variable domain of camel heavy-chain antibodies is encoded in the germline. J. Mol. Biol. 275, 413.

Nissim, A., Hoogenboom, H.R., Tomlinson, I.M., Flynn, G., Midgley, C., Lane, D., Winter, G., 1994. Antibody fragments from a 'single pot' phage display library as immunochemical reagents. EMBO J. 13, 692.

Novotny, J., Handschumacher, M., Haber, E., Bruccoleri, R., Carlson, W.B., Fanning, D.W., Smith, J.A., Rose, G.D., 1986. Antigenic determinants in proteins coincide with surface regions accessible to large probes (antibody domains). Proc. Natl. Acad. Sci. U.S.A. 83, 226.

Padlan, E.A., 1994. Anatomy of the antibody molecule. Mol. Immunol. 31, 169.

Padlan, E.A., 1996. X-ray crystallography of antibodies. Adv. Protein Chem. 49, 57.

Persson, M.A.A., Caothien, R.H., Burton, D.R., 1991. Generation of diverse high-affinity human monoclonal antibodies by repertoire cloning. Proc. Natl. Acad. Sci. U.S.A. 88, 2432.

Pessi, A., Bianchi, E., Crameri, A., Venturini, S., Tramontano, A., Sollazzo, M., 1993. A designed metal-binding protein with a novel fold. Nature 362, 367.

Riechmann, L., 1996. Rearrangement of the former VL interface in the solution structure of a camelised, single domain VH antibody. J. Mol. Biol. 259, 957.

Riechmann, L., Davies, J., 1995. Backbone assignment, secondary structure and protein A binding of an isolated, human antibody VH domain. J. Biomol. NMR 6, 141.

Riechmann, L., Cavanagh, J., McManus, S., 1991. Uniform lebeling of a recombinant antibody Fv-fragment with ^{15}N and ^{13}C for heteronuclear NMR spectroscopy. FEBS Lett. 287, 185.

Seligmann, M., Mihaesco, E., Preud'homme, J.L., Danon, F., Brouet, J.C., 1979. Heavy chain diseases: current findings and concepts. Immunol. Rev. 48, 145.

Shirai, H., Kidera, A., Nakamura, H., 1996. Structural classification of CDR-H3 in antibodies. FEBS Lett. 399, 1.

Skerra, A., 1993. Bacterial expression of immunoglobulin fragments. Curr. Opin. Immunol. 5, 256.

Skerra, A., Plückthun, A., 1988. Assembly of a functional immunoglobulin Fv fragment in *Escherichia coli*. Science 240, 1038.

Spinelli, S., Frenken, L., Bourgeois, D., de Ron, L., Bos, W., Verrips, T., Anguille, C., Cambillau, C., Tegoni, M., 1996. The crystal structure of a llama heavy chain variable domain. Nat. Struct. Biol. 3, 752.

Tomlinson, I.M., Walter, G., Jones, P.T., Dear, P.H., Sonnhammer, E.L.L., Winter, G., 1996. The imprint of somatic hypermutation on the repertoire of human germline V genes. J. Mol. Biol. 256, 813–817.

Transue, T.R., De Genst, E., Ghahroudi, M.A., Wyns, L., Muyldermans, S., 1998. Camel single domain antibody inhibits enzyme by mimicking carbohydrate substrate. Proteins: Struct., Funct., Genet. 32, 515.

Ungar-Waron, H., Elias, E., Gluckman, A., Trainin, Z., 1987. Dromedary IgG: purification, characterization and quantitation in sera of dams and newborns. Isr. J. Vet. Med. 43, 198.

Utsumi, S., Karush, F., 1964. The subunits of purified rabbit antibody. Biochemistry 3, 1329.

Vu, K.B., Ghahroudi, M.A., Wyns, L., Muyldermans, S., 1997. Comparison of llama VH sequences from conventional and heavy chain antibodies. Mol. Immunol. 34, 1121.

Ward, E.S., Güssow, D., Griffiths, A.D., Jones, P.T., Winter, G., 1989. Binding activities of a repertoire of single immunoglobulin variable domains secreted from *E. coli*. Nature 341, 544.

Webster, D.M., Henry, A.H., Rees, A.R., 1994. Antibody–antigen interactions. Curr. Opin. Struct. Biol. 4, 123.

Whitlow, M., Bell, B.A., Feng, S.-L., Filpula, D., Hardman, K.D., Hubert, S.L., Rollence, M.L., Wood, J.F., Schott, M.E.,

Milenic, D.E., Yokota, T., Schlom, J., 1993. An improved linker for scFv with reduced aggregation and enhanced proteolytic stability. Protein Eng. 6, 989.

Winter, G., Griffiths, A.D., Hawkins, R.E., Hoogenboom, H.R., 1994. Making antibodies by phage display technology. Annu. Rev. Immunol. 12, 433.

Wu, T.T., Johnson, G., Kabat, E.A., 1993. Length distribution of CDR H3 in antibodies. Proteins: Struct., Funct., Genet. 16, 1.

Yokota, T., Milenic, D.E., Whitlow, M., Schlom, J., 1990. Rapid tumor penetration of a single-chain Fv and comparison with other immunoglobulin forms. Cancer Res. 52, 3402.

US005759808A

United States Patent [19]

Casterman et al.

[11] Patent Number: 5,759,808

[45] Date of Patent: Jun. 2, 1998

[54] **IMMUNOGLOBULINS DEVOID OF LIGHT CHAINS**

[75] Inventors: **Cecile Casterman**; **Raymond Hamers**, both of Sint-Genesius-Rode, Belgium

[73] Assignee: **Vrije Universiteit Brussel**, Brussels, Belgium

[21] Appl. No.: **471,780**

[22] Filed: **Jun. 6, 1995**

Related U.S. Application Data

[62] Division of Ser. No. 106,944, Aug. 17, 1993, abandoned.

[30] **Foreign Application Priority Data**

Aug. 21, 1992 [EP] European Pat. Off. 92402326
May 21, 1993 [EP] European Pat. Off. 93401310

[51] **Int. Cl.**[6] **C12P 21/06**; C12N 5/10; C12N 1/21; C12N 15/63

[52] **U.S. Cl.** **435/69.1**; 435/243; 435/252.3; 435/320.1; 530/324; 530/325; 530/326; 530/327; 530/328; 530/387.3; 530/388.1; 536/23.53

[58] **Field of Search** 536/23.53; 435/320.1, 435/252.3, 69.1, 240.27, 243; 530/388.1, 324, 325, 326, 327, 328, 387.3

[56] **References Cited**

FOREIGN PATENT DOCUMENTS

WO 92/01787 2/1992 WIPO .

OTHER PUBLICATIONS

Shay, J. Human Hybridomas and Monoclonal Antibodies, 1985.

Webster's II New Riverside Univerity Dictionary. The Riverside Publishing Co./Houghton Mifflin Co., Boston, MA, 1995.

Hamers–Casterman et al., "Nature", 363:446–48 (1993).

Roitt et al., "Immunology", Gower Medical Publishing, London, pp. 1.5 and 5.7 (1985).

Ward et al., "Nature", 341:544–546 (1989).

Primary Examiner—Frank C. Eisenschenk
Assistant Examiner—Evelyn Rabin
Attorney, Agent, or Firm—Finnegan, Henderson, Farabow, Garrett & Dunner, L.L.P.

[57] **ABSTRACT**

There is provided an isolated immunoglobulin comprising two heavy polypeptide chains sufficient for the formation of a complete antigen binding site or several antigen binding sites, wherein the immunoglobulin is further devoid of light polypeptide chains.

8 Claims, 12 Drawing Sheets

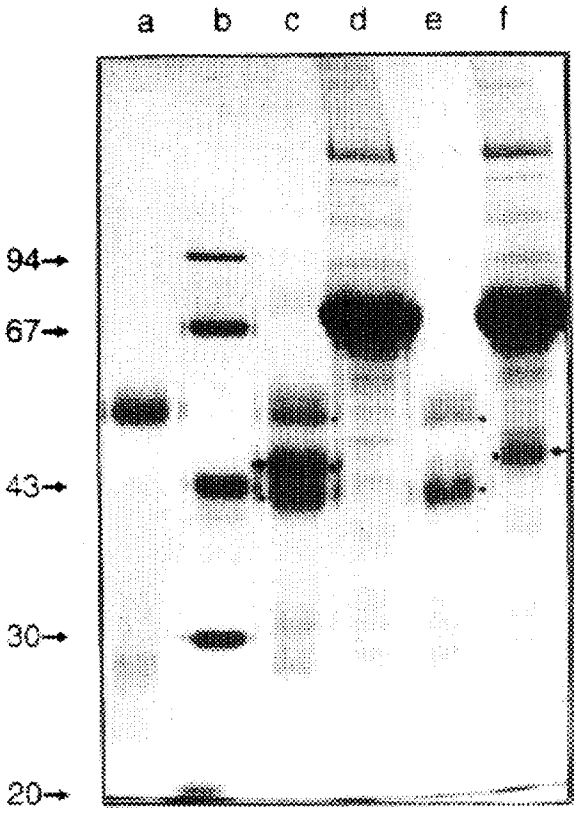


FIG. 1A

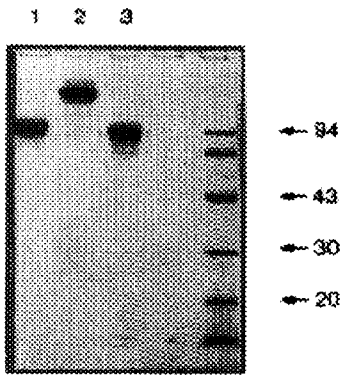


FIG. 1B

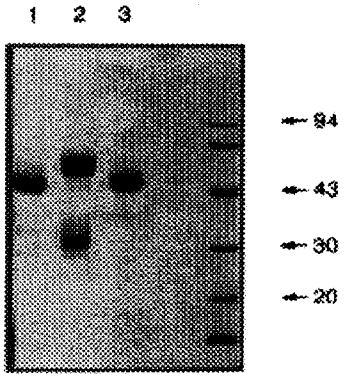


FIG. 1C

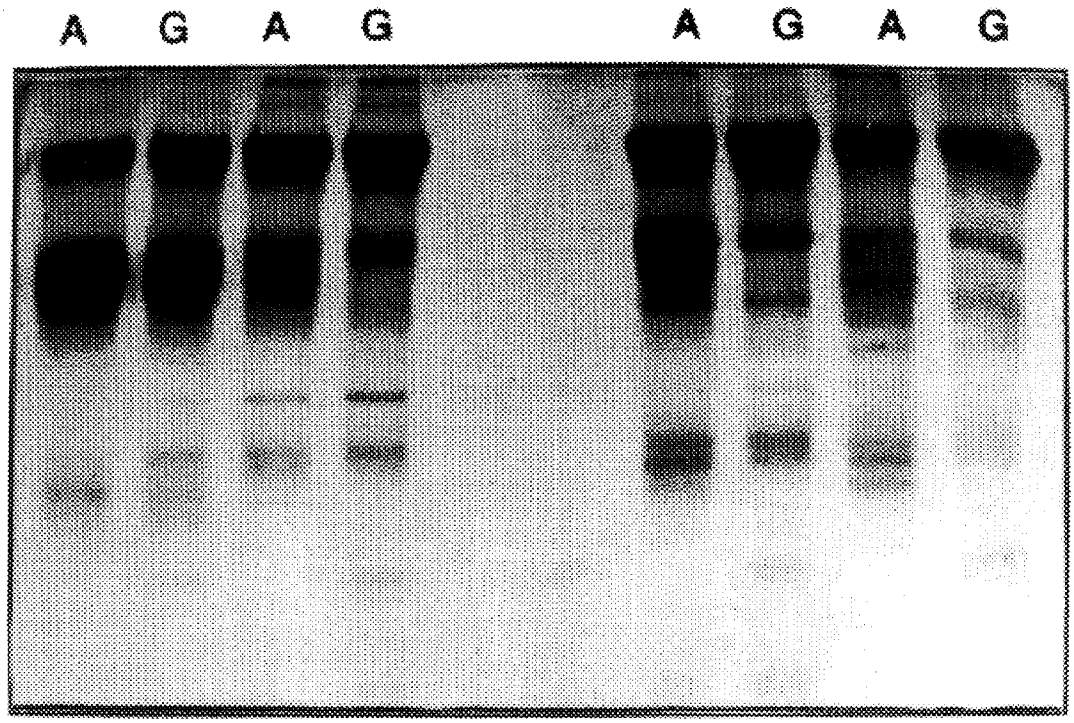

A G A G A G A G

FIG. 2A

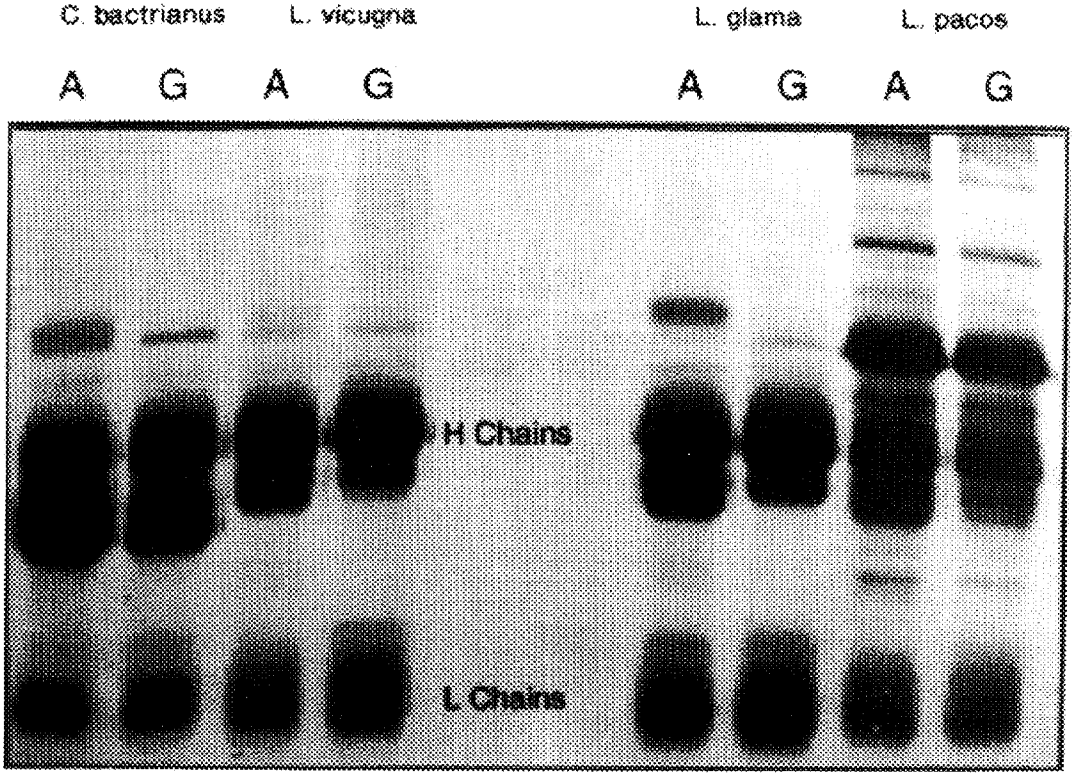
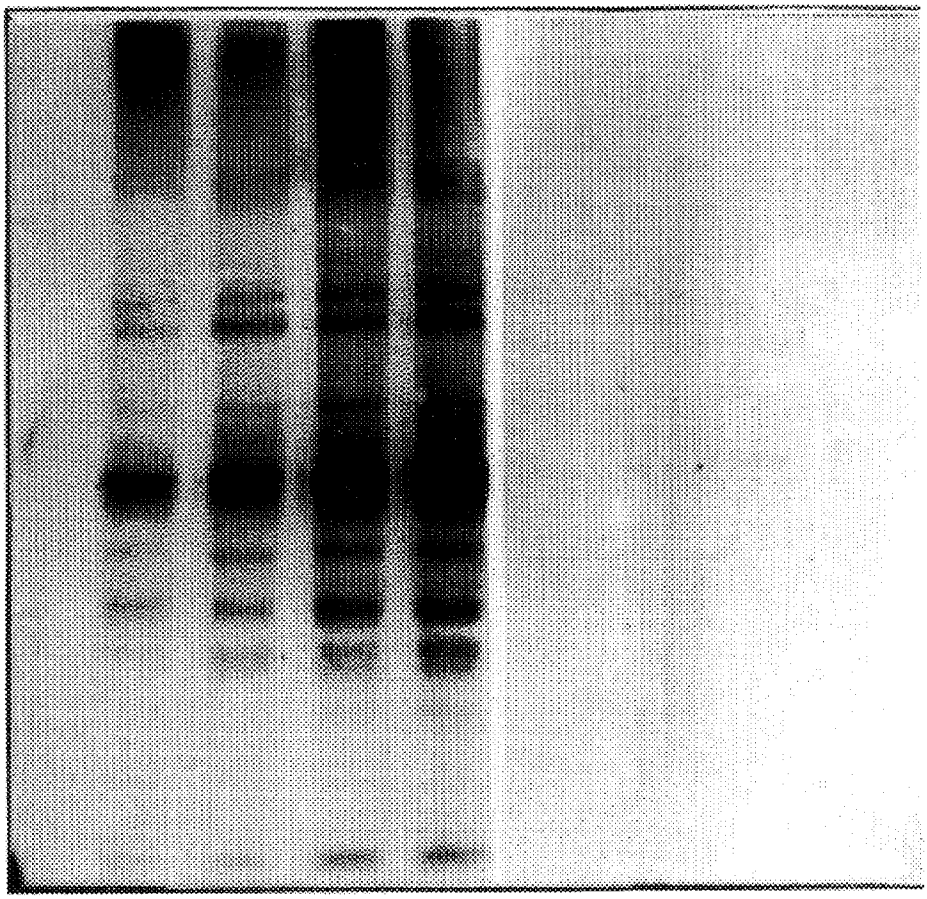

C. bactrianus
L. vicugna
L. glama
L. pacos
A G A G A G A G
H Chains
L Chains

FIG. 2B



	Prot. A	Ig1	Ig2	Ig3	Tot.Ser	Ig1	Ig2	Ig3	Tot.Ser
	Control		**T. evansi infected**				**Healthy**		
Counts/5ul	65	1258	1214	2700	2978	147	157	160	107

FIG. 3A

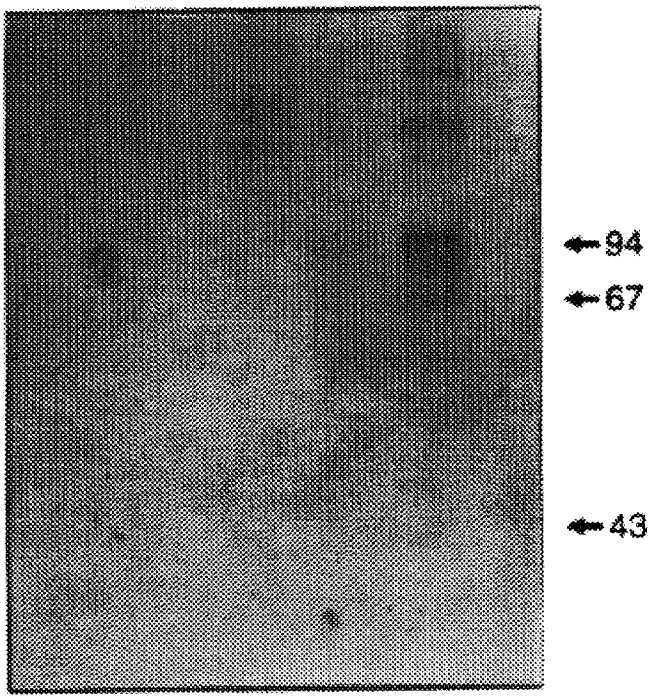

94
67
43
Ig1 Ig2 Ig3 Ig1 Ig2 Ig3
Healthy
T. evansi infected

FIG. 3B

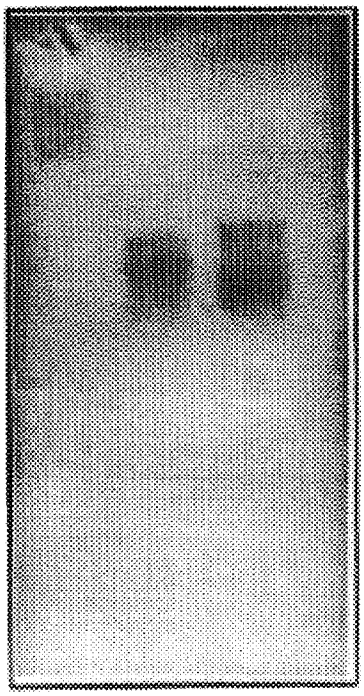

Ig1 Ig2 Ig3
T. evansi infected
Ponceau Red

FIG. 3C

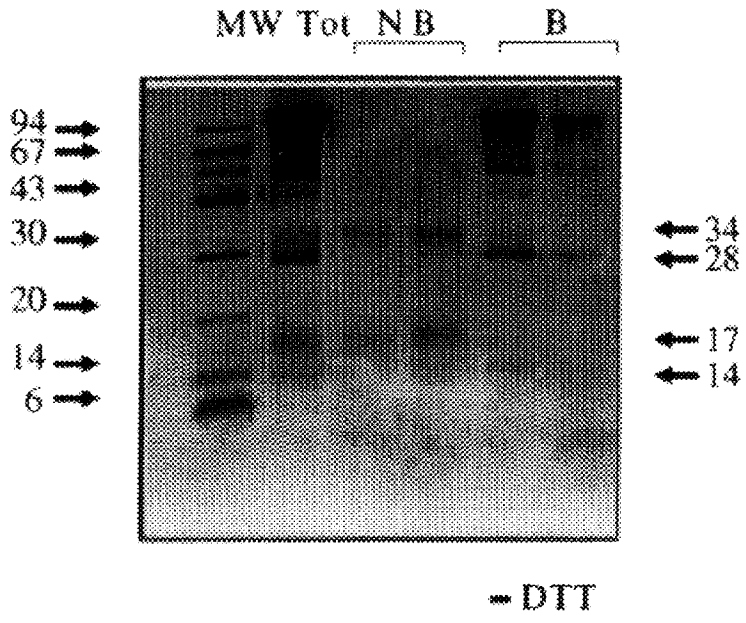

MW
Tot
N B
B
94
67
43
30
20
14
6
34
28
17
14
- DTT

FIG. 4A

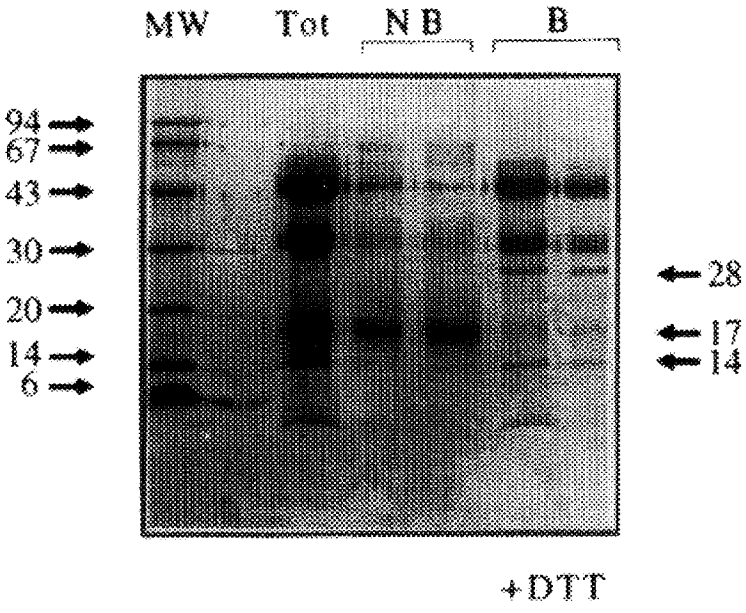

MW
Tot
N B
B
94
67
43
30
20
14
6
28
17
14
+ DTT

FIG. 4B

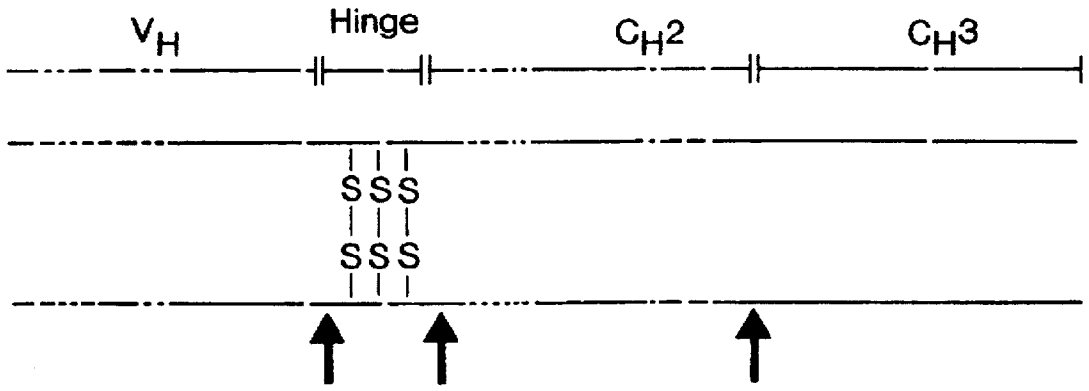


FIG. 5

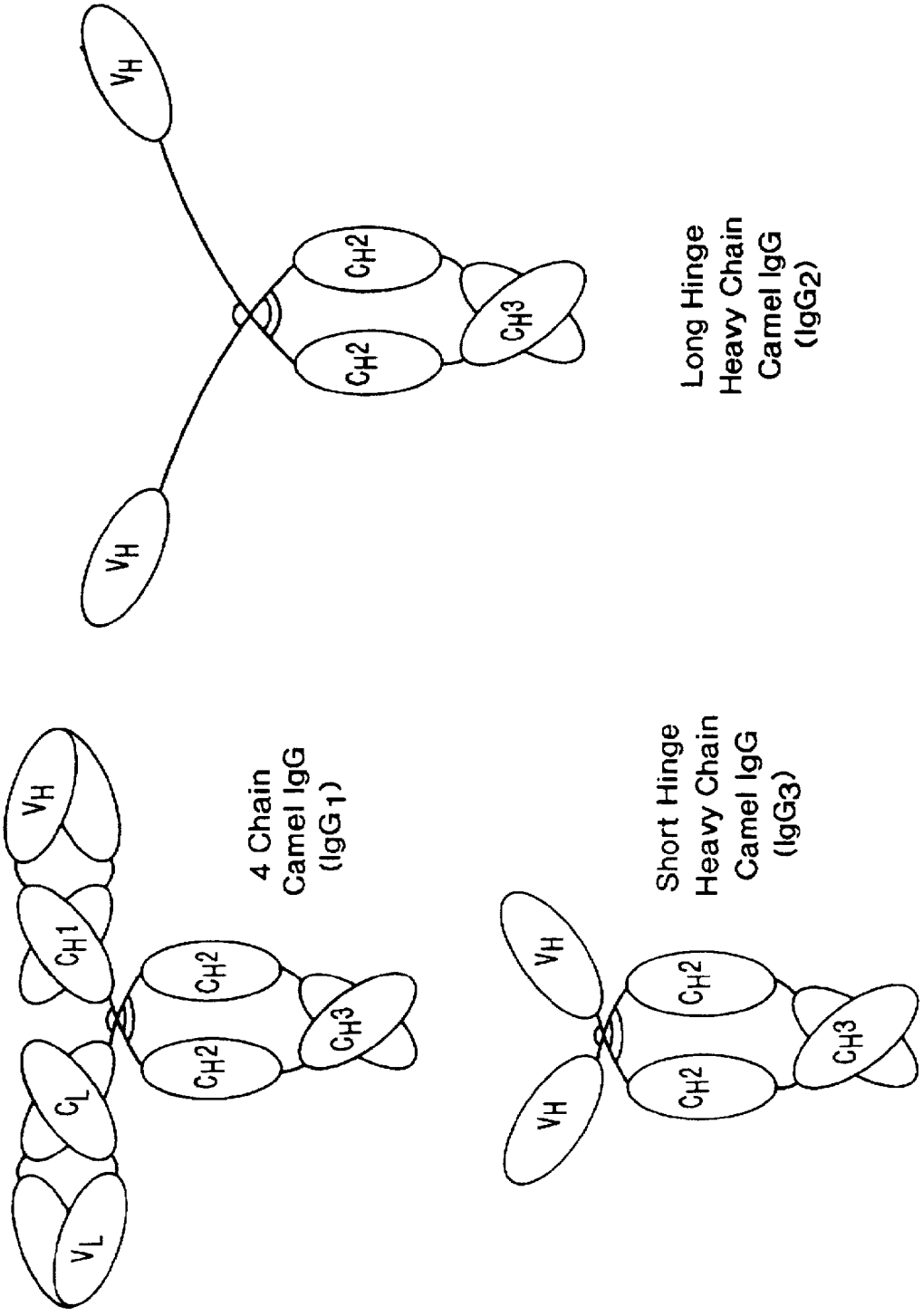

V_H
V_H
C_{H2}
C_{H2}
C_{H3}
Long Hinge Heavy Chain Camel IgG (IgG_2)
V_H
C_{H1}
C_L
V_L
C_{H2}
C_{H2}
C_{H3}
4 Chain Camel IgG (IgG_1)
V_H
V_H
C_{H2}
C_{H2}
C_{H3}
Short Hinge Heavy Chain Camel IgG (IgG_3)

FIG. 6

```
DR01006  C----------------------------TCGAG---TCTGGGGGAGG
DR27006  C----------------------------TCGAG---TCTGGGGGAGG
DR03006  C-------AGGTGA-----------AACTGCTCGAG---TCTGGAGGAGG
DR11006  C----------------------------TCGAG---TCTGGGGGAGG
DR24006  C-------AGGTGA-----------AACTGCTCGAG---TCTGGGGGAGG
DR16006  C----------------------------TCGAG---TCTGGAGGAGG
DR19006  C----------------------------TCGAG---TCTGGAGGAGG
DR07006  C----------------------------TCGAG---TCTGGGGGAGG
DR16006  C----------------------------TCGAG---TCTGGGGGAGG
DR20006  C----------------------------TCGAG---TCAGGGGGAGG
DR25006  C----------------------------TCGAG---TCTGGGGGAGG
DR20006  C----------------------------TCGAG---TCTGGAGGAGG
DR21006  C----------------------------TCGAG---TCTGGGGGAGG
DR09006  C-------AGGTGA-----------AACTGCTCGAG---TCTGGGGGAGG
DR17006  C----------------------------TCGAG---TCTGGGGGAGG
DR13006  C----------------------------TCGAG---TCAGGGGGAGG
DR02006  CTCGAGTCAGGTGTCCGGTCTGATGTGCAGCTGGTGGCGTCTGGGGGAGG

DR01006  ATCGGTGCAGGCTGGAGGGTCTCTGAGACTCTC--GTGCG-CAGCCTCTG
DR27006  CTCGGTGCAGGCTGGAGGGTCTCTGAGACTCTCCTGTGCATCTTCTTCTA
DR03006  CTCGGTGCAGACTGGAGGATCTCTGAGACTCTCCTGTGCAGT--C-TCTG
DR11006  GTCGGTGCAGGCTGGAGGGTCTCTGAGACTCTCCTGTAATGT--C-TCTG
DR24006  GTCGGTGCAGGCTGGAGGGTCTCTGAGACTCTCCTGTAATGT--C-TCTG
DR16006  CTCGGCGCAGGCTGGAGGATCTCTGAGACTCTCCTGTGCAGC--CCACGG
DR19006  CTCGGTTCAGGCTGGAGGGTCCCTTAGACTCTCCTGTGCAGC--C-TCTG
DR07006  CTCGGTGCAGGGTGGAGGGTCTCTGAGACTCTCCTGTGCAA---TCTCTG
DR16006  CTCGGTGCAGGCTGGAGGGTCTCTGAGACTCTCCTGTACAG---GCTCTG
DR20006  CTCGGTACAGGTTGGAGGGTCTCTGAGACTCTCCTGTGTAG---CCTCTA
DR25006  CTCGGTACAAACTGGAGGGTCTCTGAGACTCTCTTGCG---AAATCTCTG
DR20006  CTCGGTGCAGGCTGGAGGGTCTCTGAGACTCTCCTGTG---TAGCCTCTG
DR21006  CTCGGTGCAGGTTGGAGGGTCTCTGAAACTCTCCTGTAAAAT---CTCTG
DR09006  CTCGGTGCAGGCTGGGGGGTCTCTGACACTCTCTTGTG---TATACAC--
DR17006  CTCGGTCCAACCTGGAGGATCTCTGACACTCTCCTGTACAGTT---TCTG
DR13006  CTCGGTGGAGGCTGGAGGGTCTCTGAGACTCTCCTGTACAG---CCTCTG
DR02006  CTCGGTGCAGGCTGGAGGCTCTCTGAGACTCTCCTGTACAG---CCTCTG

DR01006  GA--TACAGTAATT---GTCCCCTCACTTG-GAGCTGGTATCGCCAGTTT
DR27006  AA--TATATGCCTT---GCACCTACGACAT-GACCTGGTACCGCCAGGCT
DR03006  GA--TTCTCCTTTA---GTACCAGTTGTAT-GGCCTGGTTCCGCCAGGCT
DR11006  GC--TCTCCCAGTA---GTACTTATTGCCT-GGGCTGGTTCCGCCAGGCT
DR24006  GC--TCTCCCAGTA---GTACTTATTGCCT-GGGCTGGTTCCGCCAGGCT
DR16006  GA--TTCCGC-TCA---ATGGTTACTACAT-CGCCTGGTTCCGTCAGGCT
DR19006  AC--TACACCATCA---CTGATTATTGCAT-GGCCTGGTTCCGCCAGGCT
DR07006  GA--TACACGTACG---GTAGCTTCTGTAT-GGGCTGGTTCCGCGAGGGT
DR16006  GA--TTCCCCTATA---GTACCTTCTGTCT-GGGGTGGTTCCGCCAGGCT
DR20006  CT--CACACCGACA---GTAGCACCTGTAT-AGGCTGGTTCCGCCAGGCT
DR25006  GA--TTGACTTTTG---ATGATTCTGACGT-GGGGTGGTACCGCCAGGCT
DR20006  GA--TTCAATTTCG---AAACTTCTCGTAT-GGCGTGGTACCGCCAGACT
DR21006  GAGGTACCCCAGATCGTGTTCCTAAATCTTTGGCCTGGTTCCGCCAGGCT
DR09006  ------CAACGATACTGGGACCA-------TGGGATGGTTTCGCCAGGCT
DR17006  --GGGCCACCTACA---GTGACTACAGTATTGGA-TGGATCCGCCAGGCT
DR13006  G----------ATACGTAT-CCT----CTATGGCCTGGTTCCGCCAGGTT
DR02006  GAGA----CAGTTTCAGTAGATT--TGCCATGTCTTGGTTCCGCCAGGCT
```

FIG. 7A

```
DR01006   CCAGGAACGGAGCGCGAGTTCGTCTCCAGTATGGATCCGGATGGAAATAC
DR27006   CCAGGCAAGGAGCGCGAATTTGTCTCAAGTATAAATATTGATGGTAAGAC
DR03006   TCAGGAAAGCAGCGTGAGGGGGTCGCAGCCATTAATAGTGGCGGTGGTAG
DR11006   CCAGGGAGGGAGCGTGAGGGGGTCACAGCGATTAA------CACTGATGG
DR24006   CCAGGGAAGGAGCGTGAGGGGGTCACAGCGATTAA------CACTGATGG
DR16006   CCTGGGAAGGGGCGTGAGGGGGTCGCAACAATTAATGGTGGTCG------
DR19006   CCAGGGAAGGAGCGTGAATTGGTCGCAGCGATTCAAGTTGTCCGTAGTGA
DR07006   CCAGGCAAGGAACGTGAGGGGATCGCAACTATTCTTAATGGTGGTACTAA
DR16006   CCAGGGAAGGAGCGTGAGGGGGTCGCGGGTATTAATAGTGCAGGAGGTAA
DR20006   CCAGGGAAGGAGCGCGAGGGGGTCGCAAGTATATATTTTGGTGATGGTGG
DR25006   CCAGGGCATGAGTGCAAATTGGTCTCAGGTATTCTGAGTGATGGTACT-C
DR20006   CCAGGAAATGTGTGTGAGTTGGTCTCAAGTATTTACAGTGATGG------
DR21006   CCAGAGAAGGAGCGCGAGGGGATCGCAGTTCTTTCGACTAAGGATGGTAA
DR09006   CCAGGGAAAGAGTGCGAAAGGGTCGCGCATATTACGCCTGATGGTATGA-
DR17006   CCAGGGAAGGACCGTGAAGTAGTCGCAGCCGCTAATACTGGTG-------
DR13006   CCAGGGCAGGAGCGCGAGGGGGTCGCGTTTGTTCAAACGG----------
DR02006   CCAGGGAAGGAGTGCGAATTGGTCTCAAGCATTCAAAGTAATGGAAGGAC

DR01006   CAAGTACA----------------CATACTCCGTGAAGGGCCGCTTCACC
DR27006   AACATACG----------------CAGACTCCGTGAAGGGCCGATTCACC
DR03006   GACATACTA-CAACACATATGTCGCCGAGTCCGTGAAGGGCCGATTCGCC
DR11006   CAGTATCAT-ATACGCA------GCCGACTCCGTGAAGGGCCGATTCACC
DR24006   CAGTGTCAT-ATACGCA------GCCGACTCCGTGAAGGGCCGATTCACC
DR16006   ------CGA-CGTCACATACTACGCCGACTCCGTGACGGGCCGATTTACC
DR19006   TACT--CGC-C-TCACAGACTACGCCGACTCCGTGAAGGGACGATTCACC
DR07006   -------------CACATACTATGCCGACTCGGTGAAGGGCCGATTCACC
DR16006   -------------TACTTACTATGCCGACGCCGTGAAGGGCCGATTCACC
DR20006   -------------TACGAATTATCGCGACTCCGTGAAGGGCCGATTCACC
DR25006   CATATACAAAGAGTGGAGACTATGCTGAGTCTGTGAGGGGCCGGGTTACC
DR20006   CA-AAACATACTACGTCGACC--GCA-------TGAAGGGCCGATTCACC
DR21006   GA-------------CATTCTATGCCGACTCCGTGAAGGGCCGATTCACC
DR09006   ---------------CCTTCATTGATGAACCCGTGAAGGGGCGATTCACG
DR17006   ------CGACTAGTAAATTCTACGTCGACTTTGTGAAGGGCCGATTCACC
DR13006   --CTGACAAT-AGTGCATTATATGGCGACTCCGTGAAGGGCCGATTCACC
DR02006   AACTGA----------------GGCCGATTCCGTGCAAGGCCGATTCACC

DR01006   ATGTCCCGAGGCAGCACCGAGTACACAGTATTTCTGCAAATGGACAATCT
DR27006   ATCTCCCAAGACAGCGCCAAGAACACGGTGTATCTGCAGATGAACAGCCT
DR03006   ATCTCCCAAGACAACGCCAAGACCACGGTATATCTTGATATGAACAACCT
DR11006   ATCTCCCAAGACACCGCCAAGGAAACGGTACATCTCCAGATGAACAACCT
DR24006   ATCTCCCAAGACACCGCCAAGAAAACGGTATATCTCCAGATGAACAACCT
DR16006   ATCTCCCGAGACAGCCCCAAGAATACGGTGTATCTGCAGATGAACAGCCT
DR19006   ATCTCCCAAGGCAACACCAAGAACACAGTGAATCTGCAAATGAACAGCCT
DR07006   ATCTCCCAAGACAGCACGTTGAAGACGATGTATCTGCTAATGAACAACCT
DR16006   ATCTCCCAAGGGAATGCCAAGAATACGGTGTTTCTGCAAATGGATAACTT
DR20006   ATCTCCCAACTCAACGCCCAGAACACAGTGTATCTGCAAATGAACAGCCT
DR25006   ATCTCCAGAGACAACGCCAAGAACATGATATACCTTCAAATGAACGACCT
DR20006   ATTTCTAGAGAGAATGCCAAGAATACATTGTATCTACAACTGAGCGGCCT
DR21006   ATCTTCTTAGATAATGACAAGACCACTTTCTCCTTACAACTTGATCGACT
DR09006   ATCTCCCGAGACAACGCCCAGAAAACGTTGTCTTTGCGAATGAATAGTCT
DR17006   ATTTCCCAAGACAACGCCAAGAATACGGTATATCTGCAAATGAGCTTCCT
DR13006   ATCTCCCACGACAACGCCAAGAACACGCTGTATCTGCAAATGCGCAACCT
DR02006   ATCTCCCGAGACAATTCCAGGAACACAGTGTATCTGCAAATGAACAGCCT
```

FIG. 7B

```
DRO1006  GAAACCTGAGGACACGGCGATGTATTACTGTAAAAC-A---GCCCTAC--
DR27006  GAAACCTGAGGACACGGCGATGTATTACTGTAAAAT-A---GA--TTC--
DRO3006  AACCCCTGAAGACACGGCTACGTATTACTGTGCGGCGG---TCCCAGCCC
DR11006  GCAACCTGAGGATACGGCCACCTATTACTGCGCGGCAA---GACTGACGG
DR24006  GCAACCTGAGGATACGGCCACCTATTACTGCGCGGCAA---GACTGACGG
DR16006  GAAACCTGAGGACACGGCCATCTACTTCTGTGCAGCAG---G-----CTC
DR19006  GACACCTGAGGACACGGCCATCTACAGTTGTGCGGCAA---C-----CAG
DRO7006  GAAACCTGAAGACACGGGCACCTATTACTGTGCTG-CA---GAACTAAGT
DR16006  GAAACCTGAGGACACGGCCATCTATTACTGCGCGG-CG---GATAGTCCA
DR20006  GAAACCTGAGGACAGCGCCATGTACTACTGTGCAATCA---CTGAAATTG
DR25006  GAAACCTGAGGACACGGCCATGTATTACTGCGCGGTAGATGGTTGGACCC
DR20006  CAAACCTGAGGACACGGCCATGTATTACTGTGCG--------------CC
DR21006  GAACCCGGAGGACACTGCCGACTACTACTGCGCTGCAAATCAATTAGC--
DRO9006  GAGGCCTGAGGACACGGCCGTGTATTACTGTGCGGCAGATTG--------
DR17006  GAAACCTGAGGACACGGCCATCTATTACTGTGCGGCAG-----CGGACCC
DR13006  GCAACCTGACGACACTGGCGTGTACTACTGTGCGGCC----------CAA
DRO2006  GAAACCCGAGGACACGGCCGTGTATTACTGTGGGGCAGT----------

DRO1006  ------------------------A-AC--CTGGGGGTTATTGTGGGTA-
DR27006  ------------------------GTAC--CCGTGCCATCTCCTTGATG-
DRO3006  ACTTGGGACCT--------------GGCG-CCATT------CTTGATTTG
DR11006  AGATGGGGGCTTGTGATGCGAGATGGGCGACCTTAGC--GACAAGGAC-G
DR24006  AGATGGGGGCTTGTGATGCGAGATGGGCGACCTTAGC--GACAAGGAC-G
DR16006  GCGTTTTT-CTAGTCCTGTTGGGAGCACTTC-TAGAC---TCGAAAGTAG
DR19006  TAGTTTTTACTGGTACT-------GCAC--------C---ACG------G
DRO7006  GGTGGTAGTTGTGAATTGC---CTTTGC------TATTTGACTA------
DR16006  TGTTACATGCCGACTATGC---CCGCTCCCCCGATACGAGACAGTTTTGG
DR20006  AGTGGTATGGGTGCAATTT---AAGGACTACTTTTACT---C-------G
DR25006  GGAAGGAAG--GGGGAATCGGGTTAC----CCTGGTCGGTCCAATGTGAA
DR20006  GGTTGAA---------------TATC----CTATTGCAGAC--ATGTGTT
DR21006  ---TGGTGGCTGGTATT------TGGACCCGAATTACTGG-CTCTCTGTG
DRO9006  ---GAAATACTGGA----CTTGTGGTGC--CCAGA-CTGG--------AG
DR17006  AAGTATATATTATAGTATC--------CTCCNNAT---------------
DR13006  AAGAAGGATCGTA-----CTAGATGGGC--------CGAGCCT-------
DRO2006  --------------------------CTCCCTAA--TGGACCGAATTTC

DRO1006  --TGGGTANTGCCTCTGGGGCCAGGGGACCCAGGTCACCGTCTCCTCACT
DR27006  --T-----------CTGGGGCCAGGGGACCCAGGTCACCGTCTCCTCACT
DRO3006  AAAAAGTATAAGTACTGGGGCCAGGGGACCCAGGTCACCGTCTCCTCACT
DR11006  TTTGCGTATAACTACTGGGGCCGGGGGACCCAGGTCACCGTCTCCTCACT
DR24006  TTTGCGTATAACTACTGGGGCCGGGGGACCCAGGTCACCGTCTCCTCACT
DR16006  CGA-CT-ATAACTATTGGGGCCAGGGGATCCAGGTCACCGTCACCTCACT
DR19006  CGC-CTTATAACGTCTGGGGTCAGGGGACCCAGGTCACCGTCTCCTCACT
DRO7006  CTGGG----------GCCAGGGCACCCAGGTCACCGTCTCCTCACT
DR16006  CTGGGATGATTTT-----GGCCAGGGGACCCAGGTCACCGTCTCCTCACT
DR20006  CTGGG----------GCCAGGGGACCCAGGTCACCGTCTCCTCACT
DR25006  GATGGTTATAACTATTGGGGCCAGGGGACCCAGGTCACCGTCTCCTCAC-
DR20006  CGAGAT----ACG---GCGACCCGGGGACCCAGGTCACCGTCTCCTCAC-
DR21006  GGTGCATATGCCATCTGGGGCCAGGGGACCCAGGTCACCGTCTCCTCAC-
DRO9006  GATACTTCGGACAG-TGGGGTCAGGGGGCCCAGGTCACCGTCTCCTCACT
DR17006  --TGAGTATAAGTACTGGGGCCAGGGGACCCAGGTCACCGTCTCCTCA--
DR13006  CGAGAATGGAACAACTGGGGCCAGGGGACCCAGGTCACCGTCTCCTCA--
DRO2006  CCAACATGGG--TGCCGGGGCCAGGGAACCCAGGTCACCGTCTCCT----
```

FIG. 7C

```
DR01006   AG----TTACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTC
DR27006   AG----TTACCCGTACGAGCTTCCGGACTACGGTTCTTAATAGAATTC
DR03006   AGCTAGTTACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTC
DR11006   AG----TTACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTC
DR24006   AGCTAGTTACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTC
DR16006   ----AGTTACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTC
DR19006   ----AGTTACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTC
DR07006   ----AGTTACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTC
DR16006   ----AGTTACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTC
DR20006   ----AGTTACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTC
DR25006   ---TAGTTACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTC
DR20006   ---TAGTTACCCGTACGACGAACCGGACTACGGTTCTTAATAGAATTC
DR21006   ---TAGTTACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTC
DR09006   AGCTAGTTACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTC
DR17006   ----------------------------------------------
DR13006   ----------------------------------------------
DR02006   ----------------------------------------------TA
```

FIG. 7D

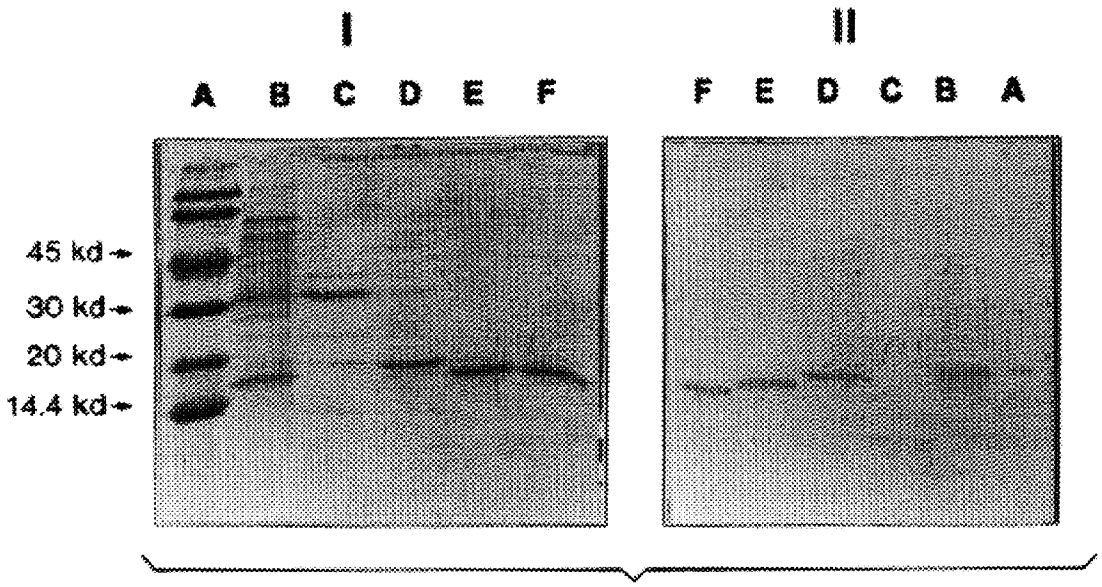


FIG. 8

IMMUNOGLOBULINS DEVOID OF LIGHT CHAINS

This is a division of application Ser. No. 08/106,944, filed Aug. 17, 1993, now abandoned.

The invention relates to new isolated immunoglobulins which are devoid of light polypeptide chains. These immunoglobulins do not consist in the degradation products of immunoglobulins composed of both heavy polypeptide and light polypeptide chains but to the contrary, the invention defines a new member of the family of the immunoglobulins, especially a new type of molecules capable of being involved in the immune recognition. Such immunoglobulins can be used for several purposes, especially for diagnosis or therapeutical purposes including protection against pathological agents or regulation of the expression or activity of proteins.

Up to now the structure proposed for immunoglobulins consists of a four-chain model referring to the presence of two identical light polypeptide chains (light chains) and two identical heavy polypeptide chains (heavy chains) linked together by disulfide bonds to form a y- or T-shaped macromolecules. These chains are composed of a constant region and a variable region, the constant region being subdivided in several domains. The two heavy polypeptide chains are usually linked by disulphide bounds in a so-called "hinge region" situated between the first and second domains of the constant region.

Among the proteins forming the class of the immunoglobulins, most of them are antibodies and accordingly present an antigen binding site or several antigen binding sites.

According to the four-chain model, the antigen binding site of an antibody is located in the variable domains of each of the heavy and light chains, and requires the association of the heavy and the light chains variable domains.

For the definition of these four-chain model immunoglobulins, reference is made to Roitt. I et al (Immunology-second-Edition Gower Medical Publishing USA, 1989). Reference is especially made to the part concerning the definition of the four-chain immunoglobulins, their polypeptidic and genetic structures, the definition of their variable and constant regions and the obtention of the fragments produced by enzymatic degradation according to well known techniques.

The inventors have surprisingly established that different molecules can be isolated from animals which naturally produce them, which molecules have functional properties of immunoglobulins these functions being in some cases related to structural elements which are distinct from those involved in the function of four-chain immunoglobulins due for instance to the absence of light chains.

The invention relates to two-chain model immunoglobulins which neither correspond to fragments obtained for instance by the degradation in particular the enzymatic degradation of a natural four-chain model immunoglobulin, nor correspond to the expression in host cells, of DNA coding for the constant or the variable region of a natural four-chain model immunoglobulin or a part of these regions, nor correspond to antibodies produced in lymphopaties for example in mice, rats or human.

E. S. Ward et al (1) have described some experiments performed on variable domains of heavy polypeptide chains (V_H) or/and light polypeptide chains (V_K/F_V) to test the ability of these variable domains, to bind specific antigens. For this purpose, a library of V_H genes was prepared from the spleen genomic DNA of mice previously immunized with these specific antigens.

Ward et al have described in their publication that V_H domains are relatively sticky, presumably due to the exposed hydrophobic surface normally capped by the V_K or V_λ domains. They consequently envisage that it should be possible to design V_H domains having improved properties and further that V_H domains with binding activities could serve as the building blocks for making variable fragments (Fv fragments) or complete antibodies.

The invention does not start from the idea that the different fragments (light and heavy chains) and the different domains of these fragments of four-chain model immunoglobulin can be modified to define new or improved antigen binding sites or a four-chain model immunoglobulin.

The inventors have determined that inmunoglobulins can have a different structure than the known four-chain model and that such different immunoglobulins offer new means for the preparation of diagnosis reagents, therapeutical agents or any other reagent for use in research or industrial purposes.

Thus the invention provides new immunoglobulins which are capable of showing functional properties of four-chain model immunoglobulins although their structure appears to be more appropriate in many circumstances for their use, their preparation and in some cases for their modification. Moreover these molecules can be considered as lead structures for the modification of other immunoglobulins. The advantages which are provided by these immunoglobulins comprise the possibility to prepare them with an increased facility.

The invention accordingly relates to immunoglobulins characterized in that they comprise two heavy polypeptide chains sufficient for the formation of a complete antigen binding site or several antigen binding sites, these immunoglobulins being further devoid of light polypeptide chains. In a particular embodiment of the invention, these immunoglobulins are further characterized by the fact that they are the product of the expression in a prokaryotic or in a eukaryotic host cell, of a DNA or of a cDNA having the sequence of an immunoglobulin devoid of light chains as obtainable from lymphocytes or other cells of Camelids.

The immunoglobulins of the invention can be obtained for example from the sequences which are described in FIG. 7.

The immunoglobulins of the invention, which are devoid of light chains are such that the variable domains of their heavy chains have properties differing from those of the four-chain immunoglobulin V_H. The variable domain of a heavy-chain immunoglobulin of the invention has no normal interaction sites with the V_L or with the C_H1 domain which do not exist in the heavy chain immunoglobulins it is hence a novel fragment in many of its properties such as solubility and position of the binding site. For clarity reasons we will call it V_{HH} in this text to distinguish it from the classical V_H of four-chain immunoglobulins.

By "a complete antigen binding site" it is meant according to the invention, a site which will alone allow the recognition and complete binding of an antigen. This could be verified by any known method regarding the testing of the binding affinity.

These immunoglobulins which can be prepared by the technique of recombinant DNA, or isolated from animals, will be sometimes called "heavy-chain immunoglobulins" in the following pages. In a preferred embodiment of the invention, these immunoglobulins are in a pure form.

In a first embodiment, the immunoglobulins of the invention are obtainable in prokaryotic cells, especially in *E. coli* cells by a process comprising the steps of:

a) cloning in a Bluecript vector of a DNA or cDNA sequence coding for the V_{HH} domain of an immunoglobulin devoid of light chain obtainable for instance from lymphocytes of Camelids,

b) recovering the cloned fragment after amplification using a 5' primer containing an Xho site and a 3' primer containing the Spe site having the following sequence TC TTA ACT ACT GAG GAG ACG GTG ACC TG, SEQ ID NO: 51

c) cloning the recovered fragment in phase in the immuno PBS vector after digestion of the vector with Xho and Spe restriction enzymes,

d) transforming host cells, especially *E. coli* by transfection with the recombinant immuno PBS vector of step c,

e) recovering the expression product of the V_{HH} coding sequence, for instance by using antibodies raised against the dromadary V_{HH} domain.

In another embodiment the immunoglobulins are heterospecific immunoglobulins obtainable by a process comprising the steps of:

obtaining a first DNA or cDNA sequence coding for a V_{HH} domain or part thereof having a determined specificity against a given antigen and comprised between Xho and Spe sites,

obtaining a second DNA or cDNA sequence coding for a V_{HH} domain or part thereof, having a determined specificity different from the specificity of the first DNA or cDNA sequence and comprised between the Spe and EcoRI sites,

digesting an immuno PBS vector with EcoRI and XhoI restriction enzymes,

ligating the obtained DNA or cDNA sequences coding for V_{HH} domains, so that the DNA or CDNA sequences are serially cloned in the vector,

transforming a host cell, especially *E. coli* cell by transfection, and recovering the obtained immunoglobulins.

In another embodiment, the immunoglobulins are obtainable by a process comprising the steps of:

obtaining a DNA or CDNA sequence coding for a V_{HH} domain or part thereof, having a determined specific antigen binding site,

amplifying the obtained DNA or cDNA, using a 5' primer containing an initiation codon and a HindIII site, and a 3' primer containing a termination codon having a XhoI site,

recombining the amplified DNA or cDNA into the HindIII (position 2650) and XhoI (position 4067) sites of a plasmid pMM984,

transfecting permissive cells especially NB-E cells with the recombinant plasmid,

recovering the obtained products.

Successful expression can be verified with antibodies directed against a region of a V_{HH} domain, especially by an ELISA assay.

According to another particular embodiment of this process, the immunoglobulins are cloned in a parvovirus.

In another example these immunoglobulins are obtainable by a process comprising the further cloning of a second DNA or cDNA sequence having another determined antigen binding site, in the pMM984 plasmid.

Such an Immunoglobulin can be further characterized in that it is obtainable by a process wherein the vector is Yep 52 and the transformed recombinant cell is a yeast especially *S. cerevisiae*.

A particular Immunoglobulin is characterized in that it has a catalytic activity, especially in that it is directed against an antigen mimicking an activiated state of a given substrate. These catalytic antibodies can be modified at the level of their biding site, by random or directed mutagenesis in order to increase or modify their catalytic function. Reference may be made to the publication of Lerner et al (TIBS November 1987. 427–430) for the general technique for the preparation of such catalytic immunoglobulins.

According to a preferred embodiment, the immunoglobulins of the invention are characterized in that their variable regions contain in position 45, an amino-acid which is different from leucine, proline or glutamine residue.

Moreover the heavy-chain immunoglobulins are not products characteristic of lymphocytes of animals nor from lymphocytes of a human patient suffering from lymphopathies. Such immunoglobulins produced in lymphopathies are monoclonal in origin and result from pathogenic mutations at the genomic level. They have apparently no antigen binding site.

The two heavy polypeptide chains of these immunoglobulins can be linked by a hinge region according to the definition of Roitt et al.

In a particular embodiment of the invention, immunoglobulins corresponding to the above-defined molecules are capable of acting as antibodies.

The antigen binding site(s) of the immunoglobulins of the invention are located in the variable region of the heavy chain.

In a particular group of these immunoglobulins each heavy polypeptide chain contains one antigen binding site on its variable region, and these sites correspond to the same amino-acid sequence.

In a further embodiment of the invention the immunoglobulins are characterized in that their heavy polypeptide chains contain a variable region (V_{HH}) and a constant region (C_H) according to the definition of Roitt et al, but are devoid of the first domain of their constant region. This first domain of the constant region is called C_H1.

These immunoglobulins having no C_H1 domain are such that the variable region of their chains is directly linked to the hinge region at the C-terminal part of the variable region.

The immunoglobulins of the type described here-above can comprise type G immunoglobulins and especially immunoglobulins which are defined as immunogloblilins of class 2 (IgG2) or immunoglobulins of class 3 (IgG3).

The absence of the light chain and of the first constant domain lead to a modification of the nomenclature of the immunoglobulin fragments obtained by enzymatic digestion, according to Roitt et al.

The terms Fc and pFc on the one hand, Fc' and pFc' on the other hand corresponding respectively to the papain and pepsin digestion fragments are maintained.

The terms Fab $F(ab)_2$ $F(ab')_2$ Fabc, Fd and Fv are no longer applicable in their original sense as these fragments have either a light chain, the variable part of the light chain or the C_H1 domain.

The fragments obtained by papain digestion and composed of the V_{HH} domain and the hinge region will be called $FV_{HH}h$ or $F(V_{HH}h)_2$ depending upon whether or not they remain linked by the disulphide bonds.

In another embodiment of the invention, immunoglobulins replying to the hereabove given definitions can be originating from animals especially from animals of the camelid family. The inventors have found out that the heavy-chain immunoglobulins which are present in camelids are not associated with a pathological situation which would induce the production of abnormal antibodies with respect to the four-chain immunoglobulins. On the basis of a comparative study of old world camelids (*Camelus bac-*

trianus and *Camelus dromaderius*) and new world camelids (for example Lama Paccos, Lama Glama, and Lama Vicuqna) the inventors have shown that the immunoglobulins of the invention, which are devoid of light polypeptide chains are found in all species. Nevertheless differences may be apparent in molecular weight of these immunoglobulins depending on the animals. Especially the molecular weight of a heavy chain contained in these immunoglobulins can be from approximately 43 kd to approximately 47 kd, in particular 45 kd.

Advantageously the heavy-chain immunoglobulins of the invention are secreted in blood of camelids.

Immunoglobulins according to this particular embodiment of the invention are obtainable by purification from serum of camelids and a process for the purification is described in details in the examples. In the case where the immunoglobulins are obtained from Camelids, the invention relates to immunoglobulins which are not in their natural biological environment.

According to the invention immunoglobulin IgG2 as obtainable by purification from the serum of camelids can be characterized in that:

it is not adsorbed by chromatography on Protein G Sepharose column,

it is adsorbed by chromatography on Protein A Sepharose column,

it has a molecular weight of around 100 kd after elution with a pH 4.5 buffer (0.15M NaCl, 0.58% acetic acid adjusted to pH 4.5 by NaOH),

it consists of heavy γ2 polypeptide chains of a molecular weight of around 46 kd preferably 45 after reduction.

According to a further embodiment of the invention another group of immunoglobulins corresponding to IgG3, as obtainable by purification from the serum of Camelids is characterized in that the immunoglobulin:

is adsorbed by chromatography on a Protein A Sepharose column,

has a molecular weight of around 100 kd after elution with a pH 3.5 buffer (0.15M NaCl, 0.58% acetic acid),

is adsorbed by chromatography on a Protein G Sepharose column and eluted with pH 3.5 buffer (0.15M NaCl, 0.58% acetic acid).

consists of heavy γ3 polypeptide chains of a molecular weight of around 45 Kd in particular between 43 and 47 kd after reduction.

The immunoglobulins of the invention which are devoid of light chains, nevertheless comprise on their heavy chains a constant region and a variable region. The constant region comprises different domains.

The variable region of immunoglobulins of the invention comprises frameworks (FW) and complementarity determining regions (CDR), especially 4 frameworks and 3 complementarity regions. It is distinguished from the four-chain immunoglobulins especially by the fact that this variable region can itself contain an antigen binding site or several, without contribution of the variable region of a light chain which is absent.

The amino-acid sequences of frameworks 1 and 4 comprise among others respectively amino-acid sequences which can be selected from the following for the framework 1 domain

G G S V Q T G G S L R L S C E I S G L T F D SEQ ID NO:1

G G S V Q T G G S L R L S C A V S G F S F S SEQ ID NO:2

G G S E Q G G G S L R L S C A I S G Y T Y G SEQ ID NO:3

G G S V Q P G G S L T L S C T V S G A T Y S SEQ ID NO:4

G G S V Q A G G S L R L S C T G S G F P Y S SEQ ID NO:5

G G S V Q A G G S L R L S C V A G F G T S SEQ ID NO:6

G G S V Q A G G S L R L S C V S F S P S S SEQ ID NO:7

for the framework 4 domain

W G Q G T Q V T V S S SEQ ID NO:8

W G Q G T L V T V S S SEQ ID NO:9

W G Q G A Q V T V S S SEQ ID NO:10

W G Q G T Q V T A S S SEQ ID NO:11

R G Q G T Q V T V S L SEQ ID NO:12

for the CDR3 domain

A L Q P G G Y C G Y G X - - - - - - - - - - C L SEQ ID NO:62

V S L M D R I S Q H - - - - - - - - - - - - G C SEQ ID NO:63

V P A H L G P G A I L D L K K Y - - - - - - K Y SEQ ID NO:64

F C Y S T A G D G G S G E - - - - - - - - - M Y SEQ ID NO:65

E L S G G S C E L P L L F - - - - - - - - - D Y SEQ ID NO:66

D W K Y W T C G A Q T G G Y F - - - - - - - G Q SEQ ID NO:67

R L T E M G A C D A R W A T L A T R T F A Y N Y SEQ ID NO:68

Q K K D R T R W A E P R E W - - - - - - - - N N SEQ ID NO:69

G S R F S S P V G S T S R L E S - S D Y - - N Y SEQ ID NO:70

A D P S I Y Y S I L X I E Y - - - - - - - - K Y SEQ ID NO:71

D S P C Y M P T M P A P P I R D S F G W - - D D SEQ ID NO:72

T S S F Y W Y C T T A P Y - - - - - - - - - N V SEQ ID NO:73

T E I E W Y G C N L R T T F - - - - - - - - T R SEQ ID NO:74

-continued

N Q L A G G W Y L D P N Y W L S V G A Y - - A I SEQ ID NO:75

R L T E M G A C D A R W A T L A T R T F A Y N Y SEQ ID NO:76

D G W T R K E G G I G L P W S V Q C E D G Y N Y SEQ ID NO:77

D S Y P C H L L - - - - - - - - - - - - - - - - D V SEQ ID NO:78

V E Y P I A D M C S - - - - - - - - - - - - - - R Y SEQ ID NO:79

As stated above, the immunoglobulins of the invention are preferably devoid of the totality of their C_H1 domain.

Such immunoglobulins comprise C_H2 and C_H3 domains in the C-terminal region with respect to the hinge region.

According to a particular embodiment of the invention the constant region of the immunoglobulins comprises C_H2 and C_H3 domains comprising an amino-acid sequence selected from the following for the C_H2 domain:

APELLGGPTVFIFPPKPKDVLSITLTP SEQ ID NO: 31
APELPGGPSVFVFPTKPKDVLSISGRP SEQ ID NO: 32
APELPGGPSVFVFPPKPKDVLSISGRP SEQ ID NO: 33
APELLGGPSVFIFPPKPKDVLSISGRP SEQ ID NO: 34

for the C_H3 domain:

GQTREPQVYTLA SEQ ID NO: 35
GQTREPQVYTLAPXRLEL SEQ ID NO: 36
GQPREPQVYTLPPSRDEL SEQ ID NO: 109
GQPREPQVYTLPPSREEM SEQ ID NO: 110
GQPREPQVYTLPPSQEEM SEQ ID NO: 111

Interestingly the inventors have shown that the hinge region of the immunoglobulins of the invention can present variable lengths. When these immunoglobulins act as antibodies, the length of the hinge region will participate to the determination of the distance separating the antigen binding sites.

Preferably an immunoglobulin according to the invention is characterized in that its hinge region comprises from 0 to 50 amino-acids.

Particular sequences of hinge region of the immunoglobulins of the invention are the following.

GTNEVCKCPKCP SEQ ID NO: 37

or,

EPKIPQPQPKPQPQPQPQPKPQPKPEPECTCPKCP SEQ ID NO: 38

The short hinge region corresponds to an IgG3 molecule and the long hinge sequence corresponds to an IgG2 molecule.

Isolated V_{HH} derived from heavy chain immunoglobulins or V_{HH} libraries corresponding to the heavy chain immunoglobulins can be distinguished from V_{HH} cloning of four-chain model immunoglobulins on the basis of sequence features characterizing heavy chain immunoglobulins.

The camel heavy—chain immunoglobulin V_{HH} region shows a number of differences with the V_{HH} regions derived from 4-chain immunoglobulins from all species examined. At the levels of the residues involved in the V_{HH}/V_L interactions, an important difference is noted at the level of position 45 (FW) which is practically always leucine in the 4-chain immunoglobulins (98%), the other amino acids at this position being proline (1%) or glutamine (1%).

In the camel heavy-chain immunoglobulin, in the sequences examined at present, leucine at position 45 is only found once. It could originate from a four-chain immunoglobulin. In the other cases, it is replaced by arginine, cysteine or glutamic acid residue. The presence of charged amino acids at this position should contribute to making the V_{HH} more soluble.

The replacement by camelid specific residues such as those of position 45 appears to be interesting for the construction of engineered V_{HH} regions derived from the V_{HH} repertoire of 4-chain immunoglobulins.

A second feature specific of the camelid V_{HH} domain is the frequent presence of a cysteine in the CDR_3 region associated with a cysteine in the CDR_1 position 31 or 33 or FW_2 region at position 45. The possibility of establishing a disulphide bond between the CDR_3 region and the rest of the variable domain would contribute to the stability and positioning of the binding site.

With the exception of a single pathogenic myeloma protein (DAW) such a disulphide bond has never been encountered in immunoglobulin V regions derived from 4 chain immunoglobulins.

The heavy-chain immunoglobulins of the invention have further the particular advantage of being not sticky. Accordingly these immunoglobulins being present in the serum, aggregate much less than isolated heavy chains of a four-chain immunoglobulins. The immunoglobulins of the invention are soluble to a concentration above 0.5 mg/ml, preferably above 1 mg/ml and more advantageously above 2 mg/ml.

These immunoglobulins further bear an extensive antigen binding repertoire and undergo affinity and specificity maturation in vivo. Accordingly they allow the isolation and the preparation of antibodies having defined specificity, regarding determined antigens.

Another interesting property of the immunoglobulins of the invention is that they can be modified and especially humanized. Especially it is possible to replace all or part of the constant region of these immunoglobulins by all or part of a constant region of a human antibody. For example the C_H2 and/or C_H3 domains of the immunoglobulin could be replaced by the C_H^2 and/or C_H3 domains of the IgG $\gamma 3$ human immunoglobulin.

In such humanized antibodies it is also possible to replace a part of the variable sequence, namely one or more of the framework residues which do not intervene in the binding site by human framework residues, or by a part of a human antibody.

Conversely features (especially peptide fragments) of heavy-chain immunoglobulin V_{HH} regions, could be introduced into the V_H or V_L regions derived from four-chain immunoglobulins with for instance the aim of achieving greater solubility of the immunoglobulins.

The invention further relates to a fragment of an immunoglobulin which has been described hereabove and especially to a fragment selected from the following group:

a fragment corresponding to one heavy polypeptide chain of an immunoglobulin devoid of light chains,

fragments obtained by enzymatic digestion of the immunoglobulins of the invention, especially those obtained by partial digestion with papain leading to the Fc fragment (constant fragment) and leading to $FV_{HH}h$ fragment (containing the antigen binding sites of the heavy chains) or its dimer $F(V_{HH}h)_2$, or a fragment obtained by further digestion with papain of the Fc fragment, leading to the pFc fragment corresponding to the C-terminal part of the Fc fragment,

homologous fragments obtained with other proteolytic enzymes,

a fragment of at least 10 preferably 20 amino acids of the variable region of the immunoglobulin, or the complete variable region, especially a fragment corresponding to the isolated V_{HH} domains or to the V_{HH} dimers linked to the hinge disulphide,

a fragment corresponding to the hinge region of the immunoglobulin, or to at least 6 amino acids of this hinge region,

a fragment of the hinge region comprising a repeated sequence of Pro-X,

a fragment corresponding to at least 10 preferably 20 amino acids of the constant region or to the complete constant region of the immunoglobulin.

The invention also relates to a fragment comprising a repeated sequence. Pro-X which repeated sequence contains at least 3 repeats of Pro-X, X being any amino-acid and preferably Gln (glutamine), Lys (lysine) or Glu (acide glutamique); a particular repeated fragment is composed of a 12-fold repeat of the sequence Pro-X.

Such a fragment can be advantageously used as a link between different types of molecules.

The amino-acids of the Pro-X sequence are chosen among any natural or non natural amino-acids.

The fragments can be obtained by enzymatic degradation of the immunoglobulins. They can also be obtained by expression in cells or organisms, of nucleotide sequence coding for the immunoglobulins, or they can be chemically synthesized.

The invention also relates to anti-idiotypes antibodies belonging to the heavy chain immunoglobulin classes. Such anti-idiotypes can be produced against human or animal idiotypes. A property of these anti-idiotypes is that they can be used as idiotypic vaccines, in particular for vaccination against glycoproteins or glycolipids and where the carbohydrate determines the epitope.

The invention also relates to anti-idiotypes capable of recognizing idiotypes of heavy-chain immunoglobulins.

Such anti-idiotype antibodies can be either syngeneic antibodies or allogenic or xenogeneic antibodies.

The invention also concerns nucleotide sequences coding for all or part of a protein which amino-acid sequence comprises a peptide sequence selected from the following:

G G S V Q T G G S L R L S C E I S G L T F D SEQ ID NO:1

G G S V Q T G G S L R L S C A V S G F S F S SEQ ID NO:2

G G S E Q G G G S L R L S C A I S G Y T Y G SEQ ID NO:3

G G S V Q P G G S L T L S C T V S G A T Y S SEQ ID NO:4

G G S V Q A G G S L R L S C T G S G F P Y S SEQ ID NO:5

G G S V Q A G G S L R L S C V A G F G T S SEQ ID NO:6

G G S V Q A G G S L R L S C V S F S P S S SEQ ID NO:7

W G Q G T Q V T V S S SEQ ID NO:8

W G Q G T L V T V S S SEQ ID NO:9

W G Q G A Q V T V S S SEQ ID NO:10

W G Q G T Q V T A S S SEQ ID NO:11

R G Q G T Q V T V S L SEQ ID NO:12

A L Q P G G Y C G Y G X - - - - - - - - - - C L SEQ ID NO:62

V S L M D R I S Q H - - - - - - - - - - - - G C SEQ ID NO:63

V P A H L G P G A I L D L K K Y - - - - - - K Y SEQ ID NO:64

F C Y S T A G D G G S G E - - - - - - - - - M Y SEQ ID NO:65

E L S G G S C E L P L L F - - - - - - - - - D Y SEQ ID NO:66

D W K Y W T C G A Q T G G Y F - - - - - - - G Q SEQ ID NO:67

R L T E M G A C D A R W A T L A T R T F A Y N Y SEQ ID NO:68

Q K K D R T R W A E P R E W - - - - - - - - N N SEQ ID NO:69

G S R F S S P V G S T S R L E S - S D Y - - N Y SEQ ID NO:70

A D P S I Y Y S I L X I E Y - - - - - - - - K Y SEQ ID NO:71

D S P C Y M P T M P A P P I R D S F G W - - D D SEQ ID NO:72

T S S F Y W Y C T T A P Y - - - - - - - - - N V SEQ ID NO:73

T E I E W Y G C N L R T T F - - - - - - - - T R SEQ ID NO:74

N Q L A G G W Y L D P N Y W L S V G A Y - - A I SEQ ID NO:75

R L T E M G A C D A R W A T L A T R T F A Y N Y SEQ ID NO:76

D G W T R K E G G I G L P W S V Q C E D G Y N Y SEQ ID NO:77

-continued

D S Y P C H L L - - - - - - - - - - - - - - - - D V SEQ ID NO:78

V E Y P I A D M C S - - - - - - - - - - - - - - - R Y SEQ ID NO:79

APELLGGPSVFVFPPKPKDVLSISGXPK SEQ ID NO:39
APELPGGPSVFVFPTKPKDVLSISGRPK SEQ ID NO:40
APELPGGPSVFVFPPKPKDVLSISGRPK SEQ ID NO:41
APELLGGPSVFIFPPKPKDVLSISGRPK SEQ ID NO:42
GQTREPQVYTLAPXRLEL SEQ ID NO:36
GQPREPQVYTLPPSRDEL SEQ ID NO:109
GQPREPQVYTLPPSREEM SEQ ID NO:110
GQPREPQVYTLPPSQEEM SEQ ID NO:111
VTVSSGTNEVCKCPKCPAPELPGGPSVFVFP SEQ ID NO:43

or,

VTVSSEPKIPQPQPKPQPQPQPQPKPQPKPEPECTCPKCPAPELLGGPSVFIFP SEQ ID NO:44
GTNEVCKCPKCP SEQ ID NO:37
APELPGGPSVFVFP SEQ ID NO:45
EPKIPQPQPKPQPQPQPQPKPQPKPEPECTCPKCP SEQ ID NO:38
APELLGGPSVFIFP SEQ ID NO:46

Such nucleotide sequences can be deduced from the amino-acid sequences taking into account the deneneracy of the genetic code. They can be synthesized or isolated from cells producing immunoglobulins of the invention.

A procedure for the obtention of such DNA sequences is described in the examples.

The invention also contemplates RNA, especially mRNA sequences corresponding to these DNA sequences, and also corresponding cDNA sequences.

The nucleotide sequences of the invention can further be used for the preparation of primers appropriate for the detection in cells or screening of DNA or cDNA libraries to isolate nucleotide sequences coding for immunoglobulins of the invention.

Such nucleotide sequences can be used for the preparation of recombinant vectors and the expression of these sequences contained in the vectors by host cells especially prokaryotic cells like bacteria or also eukaryotic cells and for example CHO cells, insect cells, simian cells like Vero cells, or any other mammalian cells. Especially the fact that the immunoglobulins of the invention are devoid of light chains permits to secrete them in eukaryotic cells since there is no need to have recourse to the step consisting in the formation of the BIP protein which is required in the four-chain immunoglobulins.

The inadequacies of the known methods for producing monoclonal antibodies or immunoglobulins by recombinant DNA technology comes from the necessity in the vast majority of cases to clone simultaneously the V_H and V_L domains corresponding to the specific binding site of 4 chain immunoglobulins. The animals and especially camelids which produce heavy-chain immunoglobulins according to the invention, and possibly other vertebrate species are capable of producing heavy-chain immunoglobulins of which the binding site is located exclusively in the V_{HH} domain. Unlike the few heavy-chain immunoglobulins produced in other species by chain separation or by direct cloning, the camelid heavy-chain immunoglobulins have undergone extensive maturation in vivo. Moreover their region has naturally evolved to function in absence of the V_L. They are therefore ideal for producing monoclonal antibodies by recombinant DNA technology. As the obtention of specific antigen binding clones does not depend on a stochastic process necessitating a very large number of recombinant cells, this allows also a much more extensive examination of the repertoire.

This can be done at the level of the non rearranged V_{HH} repertoire using DNA derived from an arbitrarily chosen tissue or cell type or at the level of the rearranged V_{HH} repertoire, using DNA obtained from B lymphocytes. More interesting however is to transcribe the mRNA from antibody producing cells and to clone the cDNA with or without prior amplification into an adequate vector. This will result in the obtention of antibodies which have already undergone affinity maturation.

The examination of a large repertoire should prove to be particularly useful in the search for antibodies with catalytic activities.

The invention thus provides libraries which can be generated in a way which includes part of the hinge sequence, the identification is simple as the hinge is directly attached to the V_{HH} domain.

These libraries can be obtained by cloning cDNA from lymphoid cells with or without prior PCR amplification. The PCR primers are located in the promoter, leader or framework sequences of the V_{HH} for the 5' primer and in the hinge, CH_2, CH_3, 3' untranslated region or polyA tail for the 3' primer. A size selection of amplified material allows the construction of a library limited to heavy chain immunoglobulins.

In a particular example, the following 3' primer in which a KpnI site has been constructed and which corresponds to amino-acids 313 to 319 (CGC CAT CAA GGT AAC AGT TGA SEQ ID NO: 47) is used in conjunction with mouse V_{HH} primers described by Sestry et al and containing a Xho site

AG GTC CAG CTG CTC GAG TCT GG SEQ ID NO:48
AG CTC CAG CTG CTC GAG TCT GG SEQ ID NO:49
AG GTC CAG CTT CTC GAG TCT GG SEQ ID NO:50
XhoI site

These primers yield a library of camelid heavy chain immunoglobulins comprising the V_{HH} region (related to mouse or human subgroup III), the hinge and a section of CH_2.

In another example, the cDNA is polyadenylated at its 5' end and the mouse specific V_{HH} primers are replaced by a poly T primer with an inbuilt XhoI site, at the level of nucleotide 12.

$CTCGAGT_{12}$.

The same 3' primer with a KpnI site is used.

This method generates a library containing all subgroups of immunoglobulins.

Part of the interest in cloning a region encompassing the hinge-CH_2 link is that in both $\gamma 2$ and $\gamma 3$, a Sac site is present immediately after the hinge. This site allows the grafting of the sequence coding for the V_{HH} and the hinge onto the Fc region of other immunoglobulins, in particular the human IgG_1 and IgG_3 which have the same amino acid sequence at

this site (Glu_{246} Leu_{247}). As an example, the invention contemplates a cDNA library composed of nucleotide sequences coding for a heavy-chain immunoglobulin, such as obtained by performing the following steps:

a) treating a sample containing lymphoid cells, especially periferal, lymphocytes, spleen cells, lymph nodes or another lyphoid tissue from a healthy animal, especially selected among the Camelids, in order to separate the lymphoid cells,

b) separating polyadenylated RNA from the other nucleic acids and components of the cells,

c) reacting the obtained RNA with a reverse transcriptase in order to obtain the corresponding cDNA,

d) contacting the cDNA of step c) with 5' primers corresponding to mouse V_H domain of four-chain immunoglobulins, which primer contains a determined restriction site, for example an XhoI site and with 3' primers corresponding to the N-terminal part of a C_H2 domain containing a KpnI site,

e) amplifying the DNA,

f) cloning the amplified sequence in a vector, especially in a bluescript vector,

g) recovering the clones hybridizing with a probe corresponding to the sequence coding for a constant domain from an isolated heavy-chain immunoglobulin.

This cloning gives rise to clones containing DNA sequences including the sequence coding for the hinge. It thus permits the characterization of the subclass of the immunoglobulin and the SacI site useful for grafting the $FV_{HH}h$ to the Fc region.

The recovery of the sequences coding for the heavy-chain immunoglobulins can also be achieved by the selection of clones containing DNA sequences having a size compatible with the lack of the C_H1 domain.

It is possible according to another embodiment of the invention, to add the following steps between steps c) and d) of the above process:

in the presence of a DNA polymerase and of deoxyribonucleotide triphosphates, contacting said cDNA with oligonucleotide degenerated primers, which sequences are capable of coding for the hinge region and N-terminal V_{HH} domain of an immunoglobulin, the primers being capable of hybridizing with the cDNA and capable of initiating the extension of a DNA sequence complementary to the cDNA used as template,

recovering the amplified DNA.

The clones can be expressed in several types of expression vectors. As an example using a commercially available vector Immuno PBS (Huse et al: Science (1989) 246, 1275), clones produced in Bluescript® according to the above described procedure, are recovered by PCR using the same XhoI containing 5' primer and a new 3' primer, corresponding to residues 113–103 in the framework of the immunoglobulins, in which an Spe site has been constructed: TC TTA ACT AGT GAG GAG ACG GTG ACC TG SEQ ID NO: 51. This procedure allows the cloning of the V_{HH} in the Xho Spe site of the Immuno PBS vector. However, the 3' end of the gene is not in phase with the identification "tag" and the stop codon of the vector. To achieve this, the construct is cut with Spe and the 4 base overhangs are filled in, using the Klenow fragment after which the vector is religated. A further refinement consists in replacing the marker ("tag") with a poly histidine so that metal purification of the cloned V_{HH} can be performed. To achieve this a Spe/EcoRI double stranded oligonucleotide coding for 6 histidines and a termination codon is first constructed by synthesis of both strands followed by heating and annealing:

CTA GTG CAC CAC CAT CAC CAT CAC TAA* TAG* SEQ ID NO:52
AC GTG GTG GTA GTG GTA GTG ATT ATC TTA A
SEQ ID NO:53

The vector containing the insert is then digested with SpeI and EcoRI to remove the resident "tag" sequence which can be replaced by the poly-His/termination sequence. The produced V_{HH} can equally be detected by using antibodies raised against the dromedary V_{HH} regions. Under laboratory conditions, V_{HH} regions are produced in the Immuno PBS vector in mg amounts per liter.

The invention also relates to a DNA library composed of nucleotide sequences coding for a heavy-chain immunoglobulin, such as obtained from cells with rearranged immunoglobulin genes.

In a preferred embodiment of the invention, the library is prepared from cells from an animal previously immunized against a determined antigen. This allows the selection of antibodies having a preselected specificity for the antigen used for immunization.

In another embodiment of the invention, the amplification of the cDNA is not performed prior to the cloning of the cDNA.

The heavy-chain of the four-chain immunoglobulins remains sequestered in the cell by a chaperon protein (BIP) until it has combined with a light chain. The binding site for the chaperon protein is the C_H1 domain. As this domain is absent from the heavy chain immunoglobulins, their secretion is independent of the presence of the BIP protein or of the light chain. Moreover the inventors have shown that the obtained immunoglobulins are not sticky and accordingly will not abnormally aggregate.

The invention also relates to a process for the preparation of a monoclonal antibody directed against a determined antigen, the antigen binding site of the antibody consisting of heavy polypeptide chains and which antibody is further devoid of light polypeptide chains, which process comprises

immortalizing lymphocytes, obtained for example from the peripheral blood of Camelids previously immunized with a determined antigen, with an immortal cell and preferably with myeloma cells, in order to form a hybridoma,

culturing the immortalized cells (hybridoma) formed and recovering the cells producing the antibodies having the desired specificity.

The preparation of antibodies can also be performed without a previous immunization of Camelids.

According to another process for the preparation of antibodies, the recourse to the technique of the hybridoma cell is not required.

According to such process, antibodies are prepared in vitro and they can be obtained by a process comprising the steps of:

cloning into vectors, especially into phages and more particularly filamentous bacteriophages, DNA or cDNA sequences obtained from lymphocytes especially PBLs of Camelids previously immunized with determined antigens,

transforming prokaryotic cells with the above vectors in conditions allowing the production of the antibodies,

selecting the antibodies for their heavy-chain structure and further by subjecting them to antigen-affinity selection,

recovering the antibodies having the desired specificity,

In another embodiment of the invention the cloning is performed in vectors, especially into plasmids coding for bacterial membrane proteins. Procaryotic cells are then transformed with the above vectors in conditions allowing the expression of antibodies in their membrane.

The positive cells are further selected by antigen affinity selection.

The heavy chain antibodies which do not contain the C_H1 domain present a distinct advantage in this respect. Indeed, the C_H1 domain binds to BIP type chaperone proteins present within eukaryotic vectors and the heavy chains are not transported out of the endocytoplasmic reticulum unless light chains are present. This means that in eukaryotic cells, efficient cloning of 4-chain immunoglobulins in non mammalian cells such as yeast cells can depend on the properties of the resident BIP type chaperone and can hence be very difficult to achieve. In this respect the heavy chain antibodies of the invention which lack the CH_1 domain present a distinctive advantage.

In a preferred embodiment of the invention the cloning can be performed in yeast either for the production of antibodies or for the modification of the metabolism of the yeast. As example, Yep 52 vector can be used. This vector has the origin of replication (ORI) 2μ of the yeast together with a selection marker Leu 2.

The cloned gene is under the control of gall promoter and accordingly is inducible by galactose. Moreover, the expression can be repressed by glucose which allows the obtention of very high concentration of cells before the induction.

The cloning between BamHI and SalI sites using the same strategy of production of genes by PCR as the one described above, allows the cloning of camelid immunoglobulin genes in *E. coli*. As example of metabolic modulation which can be obtained by antibodies and proposed for the yeast, one can site the cloning of antibodies directed against cyclins, that is proteins involved in the regulation of the cellular cycle of the yeast (TIBS 16 430 J. D. McKinney, N. Heintz 1991). Another example is the introduction by genetic engineering of an antibody directed against CD_{28}, which antibody would be inducible (for instance by gall), within the genome of the yeast. The CD_{28} is involved at the level of the initiation of cell division, and therefore the expression of antibodies against this molecule would allow an efficient control of multiplication of the cells and the optimization of methods for the production in bioreactors or by means of immobilized cells.

In yet another embodiment of the invention, the cloning vector is a plasmid or a eukaryotic virus vector and the cells to be transformed are eukaryotic cells, especially yeast cells, mammalian cells for example CHO cells or simian cells such as Vero cells, insect cells, plant cells, or protozoan cells.

For more details concerning the procedure to be applied in such a case, reference is made to the publication of Marks et al. J. Mol. Biol. 1991, 222: 581–597.

Furthermore, starting from the immunoglobulins of the invention, or from fragments thereof, new immunoglobulins or derivatives can be prepared.

Accordingly immunoglobulins replying to the above given definitions can be prepared against determined antigens. Especially the invention provides monoclonal or polyclonal antibodies devoid of light polypeptide chains or antisera containing such antibodies and directed against determined antigens and for example against antigens of pathological agents such as bacteria, viruses or parasites. As example of antigens or antigenic determinants against which antibodies could be prepared, one can cite the envelope glycoproteins of viruses or peptides thereof, such as the external envelope glycoprotein of a HIV virus, the surface antigen of the hepatitis B virus.

Immunoglobulins of the invention can also be directed against a protein, hapten, carbohydrate or nucleic acid.

Particular antibodies according to the invention are directed against the galactosylα-1-3-galactose epitope.

The immunoglobulins of the invention allow further the preparation of combined products such as the combination of the heavy-chain immunoglobulin or a fragment thereof with a toxin, an enzyme, a drug, a hormone.

As example one can prepare the combination of a heavy-chain immunoglobulin bearing an antigen binding site recognizing a myeloma immunoglobulin epitope with the abrin or mistletoe lectin toxin. Such a construct would have its uses in patient specific therapy.

Another advantageous combination is that one can prepare between a heavy-chain immunoglobulins recognizing an insect gut antigen with a toxin specific for insects such as the toxins of the different serotypes of *Bacillus thuringiensis* or *Bacillus sphaericus*. Such a construct cloned into plants can be used to increase the specificity or the host range of existing bacterial toxins.

The invention also proposes antibodies having different specificities on each heavy polypeptide chains. These multifunctional, especially bifunctional antibodies could be prepared by combining two heavy chains of immunoglobulins of the invention or one heavy chain of an immunoglobulin of the invention with a fragment of a four-chain model immunoglobulin.

The invention also provides hetero-specific antibodies which can be used for the targetting of drugs or any biological substance like hormones. In particular they can be used to selectively target hormones or cytokines to a limited category of cells. Examples are a combination of a murine or human antibody raised against interleukin 2 (IL_2) and a heavy-chain antibody raised against CD_4 cells. This could be used to reactivate CD_4 cells which have lost their IL_2 receptor.

The heavy-chain immunoglobulins of the invention can also be used for the preparation of hetero-specific antibodies. These can be achieved either according to the above described method by reduction of the bridges between the different chains and reoxydation, according to the usual techniques, of two antibodies having different specificities, but it can also be achieved by serial cloning of two antibodies for instance in the Immuno pBS vector.

In such a case, a first gene corresponding to the V_{HH} domain comprised between Xho site and a Spe site is prepared as described above. A second gene is then prepared through an analogous way by using as 5' extremity a primer containing a Spe site, and as 3' extremity a primer containing a termination codon and an EcoRI site. The vector is then digested with EcoRI and XhoI and further both V_{HH} genes are digested respectively by Xo Spe and by SpeEcoRI.

After ligation, both immunoglobulin genes are serially cloned. The spacing between both genes can be increased by the introduction of addition codons within the 5' SpeI primer.

In a particular embodiment of the invention, the hinge region of IgG2 immunoglobulins according to the invention is semi-rigid and is thus appropriate for coupling proteins. In such an application proteins or peptides can be linked to various substances, especially to ligands through the hinge region used as spacer. Advantageously the fragment comprises at least 6 amino acids.

According to the invention it is interesting to use a sequence comprising a repeated sequence Pro-X, X being any amino-acid and preferably Gln, Lys or Glu, especially a fragment composed of at least a 3-fold repeat and preferably of a 12-fold repeat, for coupling proteins to ligand, or for assembling different protein domains.

The hinge region or a fragment thereof can also be used for coupling proteins to ligands or for assembling different protein domains.

Usual techniques for the coupling are appropriate and especially reference may be made to the technique of protein engineering by assembling cloned sequences.

The antibodies according to this invention could be used as reagents for the diagnosis in vitro or by imaging techniques. The immunoglobulins of the invention could be labelled with radio-isotopes, chemical or enzymatic markers or chemiluminescent markers.

As example and especially in the case of detection or observation with the immunoglobulins by imaging techniques, a label like technetium, especially technitium 99 is advantageous. This label can be used for direct labelling by a coupling procedure with the immunoglobulins or fragments thereof or for indirect labelling after a step of preparation of a complex with the technitium.

Other interesting radioactive labels are for instance indium and especially indium 111, or iodine, especially I^{131}, I^{125} and I^{123}.

For the description of these techniques reference is made to the FR patent application published under number 2649488.

In these applications the small size of the V_{HH} fragment is a definitive advantage for penetration into tissue.

The invention also concerns monoclonal antibodies reacting with anti-idiotypes of the above-described antibodies.

The invention also concerns cells or organisms in which heavy-chain immunoglobulins have been cloned. Such cells or organisms can be used for the purpose of producing heavy-chain immunoglobulins having a desired preselected specificity, or corresponding to a particular repertoire. They can also be produced for the purpose of modifying the metabolism of the cell which expresses them. In the case of modification of the metabolism of cells transformed with the sequences coding for heavy-chain immunoglobulins, these produced heavy-chain immunoglobulins are used like antisense DNA. Antisense DNA is usually involved in blocking the expression of certain genes such as for instance the variable surface antigen of trypanosomes or other pathogens. Likewise, the production or the activity of certain proteins or enzymes could be inhibited by expressing antibodies against this protein or enzyme within the same cell.

The invention also relates to a modified 4-chain immunoglobulin or fragments thereof, the V_H regions of which has been partialy replaced by specific sequences or amino acids of heavy chain immunoglobulins, especially by sequences of the V_{HH} domain. A particular modified V_H domain of a four-chain immunoglobulin, is characterized in that the leucine, proline or glutamine in position 45 of the V_H regions has been replaced by other amino acids and preferably by arginine, glutamic acid or cysteine.

A further modified V_H or V_L domain of a four-chain immunoglobulin, is characterized by linking of CDR loops together or to FW regions by the introduction of paired cysteines, the CDR region being selected between the CDR, and the CDR_3, the FW region being the FW_2 region, and especially in which one of the cysteines introduced is in position 31, 33 of the CDR, or 45 of FW_2 and the other in CDR_3.

Especially the introduction of paired cysteines is such that the CDR_3 loop is linked to the FW2 or CDR1 domain and more especially the cysteine of the CDR3 of the V_H is linked to a cysteine in position 31, 33 of the CDR, or in position 45 of FW_2.

In another embodiment of the invention, plant cells can be modified by the heavy-chain immunoglobulins according to the invention, in order that they acquire new properties or increased properties.

The heavy-chain immunoglobulins of the invention can be used for gene therapy of cancer for instance by using antibodies directed against proteins present on the tumor cells.

In such a case, the expression of one or two V_{HH} genes can be obtained by using vectors derived from parvo or adeno viruses. The parvo viruses are characterized by the fact that they are devoid of pathogenicity or almost not pathogenic for normal human cells and by the fact that they are capable of easily multiplying in cancer cells (Russel S. J. 1990, Immunol. Today II. 196–200).

The heavy-chain immunoglobulins are for instance cloned within HindIII/XbaI sites of the infectious plasmid of the murine MVM virus (pMM984). (Merchlinsky et al, 1983, J. Virol. 47, 227–232) and then placed under the control of the MVM38 promoter.

The gene of the V_{HH} domain is amplified by PCR by using a 5' primer containing an initiation codon and a HindIII site, the 3' primer containing a termination codon and a XbaI site.

This construct is then inserted between positions 2650 (HindIII) and 4067 (XbaI) within the plasmid.

The efficiency of the cloning can be checked by transfection. The vector containing the antibody is then introduced in permissive cells (NB-E) by transfection.

The cells are recovered after two days and the presence of V_{HH} regions is determined with an ELISA assay by using rabbit antiserum reacting with the V_{HH} part.

The invention further allows the preparation of catalytic antibodies through different ways. The production of antibodies directed against components mimicking activated states of substrates (as example vanadate as component mimicking the activated state of phosphate in order to produce their phosphoesterase activities, phosphonate as compound mimicking the peptidic binding in order to produce proteases) permits to obtain antibodies having a catalytic function. Another way to obtain such antibodies consists in performing a random mutagenesis in clones of antibodies for example by PCR, in introducing abnormal bases during the amplification of clones. These amplified fragments obtained by PCR are then introduced within an appropriate vector for cloning. Their expression at the surface of the bacteria permits the detection by the substrate of clones having the enzymatic activity. These two approaches can of course be combined. Finally, on the basis of the data available on the structure, for example the data obtained by XRay crystallography or NMR, the modifications can be directed. These modifications can be performed by usual techniques of genetic engineering or by complete synthesis. One advantage of the V_{HH} of the heavy chain immunoglobulins of the invention is the fact that they are sufficiently soluble.

The heavy chain immunoglobulins of the invention can further be produced in plant cells, especially in transgenics plants. As example the heavy chain immunoglobulins can be produced in plants using the pMon530 plasmid (Roger et al. Meth Enzym 153 1566 1987) constitutive plant expression vector as has been described for classical four chain antibodies (Hiat et al. Nature 342 76–78, 1989) once again using the appropriate PCR primers as described above, to generate a DNA fragment in the right phase.

Other advantages and characteristics of the invention will become apparent in the examples and figures which follow.

FIGURES

FIG. 1 Characterisation and purification of camel IgG by affinity chromatography on Protein A and Protein G sepharose (Pharmacia).

(A) shows, after reduction, the SDS-PAGE protein profile of the adsorbed and non adsorbed fractions of *Camelus dromedarius* serum. The fraction adsorbed on Protein A and eluted with NaCl 0.15M acetic acid 0.58% show upon reduction (lane c) three heavy chain components of respectively 50, 46 and 43 Kd and light chain (rabbit IgG in lane a). The fractions adsorbed on a Protein G Sepharose (Pharmacia) derivative which has been engineered to delete the albumin binding region (lane e) and eluted with 0.1M gly HCl pH 2.7 lacks the 46 Kd heavy chain which is recovered in the non adsorbed fraction (lane f). None of these components are present in the fraction non adsorbed on Protein A (lane d), lane b contains the molecular weight markers. (B) and (C) By differential elution, immunoglobulin fractions containing the 50 and 43 Kd heavy chain can be separated. 5 ml of *C. dromadarius* serum is adsorbed onto a 5 ml Protein G sepharose column and the column is extensively washed with 20 mM phosphate buffer, pH 7.0. Upon elution with pH 3.5 buffer (0.15M NaCl, 0.58% acetic acid) a 100 Kd component is eluted which upon reduction yields a 43 Kd heavy chain, (lane 1). After column eluant absorbance has fallen to background level a second immunoglobulin component of 170 Kd can be eluted with pH 2.7 buffer (0.1M glycine HC). This fraction upon reduction yields a 50 Kd heavy chain and a board light chain band (lane 2). The fraction non adsorbed on Protein G is then brought on a 5 ml Protein A Sepharose column. After washing and elution with pH 3.5 buffer (0.15M NaCl, 0.58% acetic acid) a third immunoglobulin of 100 Kd is obtained which consists solely of 46 Kd heavy chains (lane 3).

FIG. 2: Immunoglobulins of *Camelus bactrianus, Lama vicugna, Lama glama* and *Lama pacos* to Protein A (A lanes) and to Protein G (G lanes) analyzed on SDS-PAGE before (A) and after reduction (B).

10 µl of serum obtained from the different species wera added to Eppendorf® tubes containing 10 mg of Protein A or Protein G sepharose suspended in 400 µl of pH 8.3 immunoprecipitation buffer (NaCl 0.2.M, Tris 0.01M; EDTA 0.01M, Triton X100 1%, ovalbumin 0.1%). The tubes were slowly rotated for 2 hours at 4° C. After centrifugation the pellets were washed 3 times in buffer and once in buffer in which the Triton and ovalbumin had been ommitted. The pellets were then resuspended in the SDS-PAGE sample solution 70 µl per pellet with or without dithiotreitol as reductant. After boiling for 3 min at 100° C., the tubes were centrifuged and the supernatants analysed.

In all species examined the unreduced fractions (A) contain in addition to molecules of approximately 170 Kd also smaller major components of approximately 100 Kd. In the reduced sample (B) the constituant heavy and light chains are detected. In all species a heavy chain component (marked by an asterisk *) is present in the material eluted from the Protein A but absent in the material eluted from the Protein G.

FIG. 3 IgG_1, IgG_2 and IgG_3 were prepared from serum obtained from healthy or *Trypanosama evansi* infected *Camelus dromedarius* (CATT titer 1/160 (3) and analysed by radioimmunopreci- pitation or Western Blotting for anti trypanosome activity.

(A) ^{35}S methionine labelled *Trypanosome evansi* antigens lysate (500.000 counts) was added to Eppendorf tubes containing 10 µl of serum or, 20 µg of IgG_1, IgG_2 or IgG_3 in 200 µl of pH 8.3 immunoprecipitation buffer containing 0.1M TLCK as proteinase inhibitor and slowly rotated at 4° C. during one hour. The tubes were then supplemented with 10 mg of Protein A Sepharose suspended in 200 µl of the same pH 8.3 buffer and incubated at 4° C. for an additional hour.

After washing and centrifugation at 15000 rpm for 12 s, each pellet was resuspended in 75 µl SDS-PAGE sample solution containing DTT and heated for 3 min. at 100° C. After centrifugation in an Eppendorf minifuge at 15000 rpm for 30 s, 5 µl of the supernatant was saved for radioactivity determination and the reminder analysed by SDS-PAGE and fluorography. The counts/5 µl sample are inscribed on for each line.

(B) 20 µg of IgG_1, IgG_2 and IgG_3 from healthy and trypanosome infected animals were separated by SDS-PAGE without prior reduction or heating. The separated samples were then electro transferred to a nitrocellulose membrane, one part of the membrane was stained with Ponceau Red to localise the protein material and the reminder incubated with 1% ovalbumin in TST buffer (Tris 10 mM, NaCl 150 mM, Tween 0.05%) to block protein binding sites.

After blocking, the membrane was extensively washed with TST buffer and incubated for 2 hours with ^{35}S-labelled trypanosome antigen. After extensive washing, the membrane was dried and analysed by autoradiography. To avoid background and unspecific binding, the labelled trypanosome lysate was filtered through a 45µ millipore filter and incubated with healthy camel immunoglobulin and ovalbumin adsorbed on a nitrocellulose membrane.

FIG. 4: Purified IgG3 of the camel, by affinity chromatography on Protein A Sepharose are partially digested with papain and separated on Protein A sepharose.

14 mg of purified IgG3 were dissolved in 0.1M phosphate buffer pH 7.0 containing 2mM EDTA. They were digested by 1 hour incubation at 37° C. with mercurypapain (1% enzyme to protein ratio) activated by 5.10^4M cysteine. The digestion was blocked by the addition of excess iodoacetamide (4.10^2M)(13). After centrifugation of the digest in an ependorf centrifuge for 5 min at 15000 rpm, the papain fragments were separated on a protein A Sepharose column into binding (B) and non binding (NB) fractions. The binding fraction was eluted from the column with 0.1M glycine HCl buffer pH 1.7.

FIG. 5: Schematic presentation of a model for IgG3 molecules devoid of light chains.

FIG. 6:
Schematic representation of immunoglobulins having heavy polypeptide chains and devoid of light chains, regarding conventional four-chain model immunoglobulin.
Representation of a hinge region.

FIG. 7: Alignment of 17 V_{HH} DNA sequences of Camel heavy chain immunoglobulins SEQ ID NO: 92–108

FIG. 8: Expression and purification of the camel V_{HH}21 protein from *E. coli*.

I HEAVY CHAIN ANTIBODIES IN CAMELIDS

When *Camelus dromedarius* serum is adsorbed on Protein G sepharose, an appreciable amount (25–35%) of immunoglobulins (Ig) remains in solution which can then be recovered by affinity chromatography on Protein A sepharose (FIG. 1A). The fraction adsorbed on Protein G can be differentially eluted into a tightly bound fraction (25%) consisting of molecules of an unreduced apparent molecular weight (MW) of 170 Kd and a more weakly bound fraction (30–45%) having an apparent molecular weight of 100 Kd (FIG. 1B). The 170 Kd component when reduced yields 50 Kd heavy chains and large 30 Kd light chains. The 100 Kd fraction is totally devoid of light chains and appears to be solely composed of heavy chains which after reduction have on apparent MW of 43 Kd (FIG. 1C). The fraction which does not bind to Protein G can be affinity purified and eluted

from a Protein A column as a second 100 Kd component which after reduction appears to be composed solely of 46 Kd heavy chains.

The heavy chain immoglobulins devoid of light chains total up to 75% of the molecules binding to Protein A.

As all three immunoglobulins bind to Protein A we refer to them as IgG: namely IgG_1 (light chain and heavy chain γ1 (50 Kd) binding to Protein G. IgG_2 (heavy chain γ2 (46 Kd) non binding to Protein G and IgG_3 (heavy chain γ3 (43 Kd) binding to Protein G. There is a possibility that these three sub(classes) can be further subdivided.

A comparative study of old world camelids (*Camelus bactrianus* and *Camelus dromedarius*) and new world camelids (*lama pacos, lama glama, lama vicugna*) showed that heavy chain immunoglobulins are found in all species examined, albeit with minor differences in apparent molecular weight and proportion. The new world camelids differs from the old world carelids in having a larger IgG_3 molecule (heavy chain immunoglobulin binding to Protein G) in which the constituant heavy chains have an apparent molecular weight of 47 Kd (FIG. 2).

The abundance of the heavy chain immunoglobulins in the serum of camelids raises the question of what their role is in the immune response and in particular whether they bear antigen binding specificity and if so how extensive is the repertoire. This question could be answered by examining the immunoglobulins from *Trypanosoma evansi* infected camels (*Camelus dromedarius*).

For this purpose, the corresponding fractions of IgG_1, IgG_2, IgG_3 were prepared from the serum of a healthy camel and from the serum of camels with a high antitrypanosome titer, measured by the Card Agglutination Test (3). In radioimmunoprecipitation, IgG_1, IgG_2 and IgG_3 derived from infected camel indicating extensive repertoire heterogeneity and complexity (FIG. 3A) were shown to bind a large number of antigens present in a ^{35}S methionine labelled trypanosome lysate.

In blotting experiments ^{35}S methionine labelled trypanosome lysate binds to SDS PAGE separated IgG_1, IgG_2 and IgG_3 obtained from infected animals (FIG. 3B).

This leads us to conclude that the camelid heavy chain IgG_2 and IgG_3 are bona fide antigen binding antibodies.

An immunological paradigm states that an extensive antibody repertoire is generated by the combination of the light and heavy chain variable V region repertoires (6). The heavy chain immunoglobulins of the camel seem to contradict this paradigm.

Immunoglobulins are characterized by a complex I.E.F. (isoelectric focussing) pattern reflecting their extreme heterogeneity. To determine whether the two heavy chains constituting the IgG_2 and IgG_3 are identical or not, the isoelectric focussing (I.E.F.) pattern were observed before and after chain separation by reduction and alkylation using iodoacetamide as alkylating agent.

As this alkylating agent does not introduce additional charges in the molecule, the monomers resulting from the reduction and alkylation of a heavy chain homodimer will have practically the same isolectric point as the dimer, whereas if they are derived from a heavy chain heterodimer, the monomers will in most cases differ sufficiently in isoelectric point to generate a different pattern in I.E.F.

Upon reduction, and alkylation by iodoacetamide the observed pattern is not modified for the *Camelus dromedarius* IgG_2 and IgG_3 indicating that these molecules are each composed of two identical heavy chains which migrate to the same position as the unreduced molecule they originated from.

In contrast, the I.E.F. pattern of IgG_1 is completely modified after reduction as the isoelectric point of each molecule is determined by the combination of the isoelectric points of the light and heavy chains which after separation will each migrate to a different position.

These findings indicate that the heavy chains alone can generate an extensive repertoire and question the contribution of the light chain to the useful antibody repertoire. If this necessity be negated, what other role does the light chain play.

Normally, isolated heavy chain from mammalian immunoglobulins tend to aggregate considerably but are only solubilized by light chains (8, 9) which bind to the C_H1 domain of the heavy chain.

In humans and in mice a number of spontaneous or induced myelomas produce a pathological immunoglobulin solely composed of heavy chains (heavy chain disease). These myeloma protein heavy chains carry deletions in the C_H1 and V_{HH} domains (10). The reason why full length heavy chains do not give rise to secreted heavy chain in such pathological immunoglobulins seems to stem from the fact that the synthesis of Ig involves a chaperoning protein, the immunoglobulin heavy chain binding protein or BIP (11), which normally is replaced by the light chain (12). It is possible that the primordial role of the light chain in the four-chain model immunoglobulins is that of a committed heavy chain chaperon and that the emergence of light chain repertoires has just been an evolutionary bonus.

The camelid γ2 and γ3 chains are considerably shorter than the normal mammalian γ chain. This would suggest that deletions have occurred in the C_H1 domain. Differences in sizes of the γ2 and γ3 immunoglobulins of old and new world camelids suggests that deletions occurred in several evolutionary steps especially in the C_H1 domain.

II THE HEAVY CHAIN IMMUNOGLOBULINS OF THE CAMELIDS LACK THE C_H1 DOMAIN

The strategy followed for investigating the heavy chain immunoglobulin primary structure is a combination of protein and cDNA sequencing; the protein sequencing is necessary to identify sequence stretches characteristic of each immunoglobulin. The N-terminal of the immunoglobulin being derived from the heavy chain variable region repertoire only yields information on the V_{HH} subgroups (variable region of the heavy chain) and cannot be used for class or subclass identification. This means that sequence data had to be obtained from internal enzymatic or chemical cleavage sites.

A combination of papain digestion and Protein A affinity chromatography allowed the separation of various fragments yielding information on the general structure of IgG_3.

The IgG3 of the camel (*Camelus dromedarius*) purified by affinity chromatography on Protein A Sepharose were partially digested with papain and the digest was separated on Protein A Sepharose into binding and non binding fractions. These fractions were analysed by SDS PAGE under reducing and non reducing conditions (FIG. 4).

The bound fraction contained two components, one of 28 Kd and one of 14.4 Kd, in addition to uncleaved or partially cleaved material. They were well separated by gel electrophoresis (from preparative 19% SDS-PAGE gels) under non reducing conditions and were further purified by electroelution (in 50 nM amonium bicarbonate, 0.1% (w/v) SDS using a BioRad electro-eluter). After lyophilization of these electroeluted fractions, the remaining SDS was eliminated by precipitating the protein by the addition of 90% ethanol,

mixing and incubating the mixture overnight at −20° C. (14). The precipitated protein was collected in a pellet by centrifuging (15000 rpm, 5 min) and was used for protein sequencing. N-terminal sequencing was performed using the automated Edman chemistry of an Applied Biosystem 477A pulsed liquid protein sequencer. Amino acids were identified as their phenylthiohydantoin (PTH) derivatives using an Applied Biosystem 120 PTH analyser. All chemical and reagents were purchased from Applied Biosystems. Analysis of the chromatographic data was performed using Applied Biosystems software version 1.61. In every case the computer aided sequence analysis was confirmed by direct inspection of the chromatograms from the PTH analyser. Samples for protein sequencing were dissolved in either 50% (v/v) trifluoroacetic acid (TFA) (28 Kd fragment) or 100% TFA (14 Kd fragment). Samples of dissolved protein equivalent to 2000 pmol (28 Kd fragment) or 500 pmol (14 Kd fragment) were applied to TFA-treated glass fibre discs. The glass fibre discs were coated with BioBrene (3 mg) and precycled once before use.

N-terminal sequencing of the 28 Kd fragment yields a sequence homologous to the N-terminal part of γ C_H2 domain and hence to the N-terminal end of the Fc fragment. The N-terminal sequence of the 14.4 Kd fragment corresponds to the last lysine of a γ C_H2 and the N-terminal end of a γ C_H3 domain (Table 1). The molecular weight (MW) of the papain fragments and the identification of their N-terminal sequences led us to conclude that the C_H2 and C_H3 domains of the γ^3 heavy chains are normal in size and that the deletion must occur either in the C_H1 or in the V_{HH} domain to generate the shorted γ3 chain. The fractions which do not bind to Protein A Sepharose contain two bands of 34 and 17 Kd which are more diffuse is SDS PAGE indicating that they originate from the variable N-terminal part of the molecule (FIG. 4).

Upon reduction, a single diffuse band of 17 Kd is found indicating that the 34 Kd is a disulfide bonded dimer of the 17 Kd component. The 34 Kd fragment apparently contains the hinge and the N-terminal domain V_{HH}.

The protein sequence data can be used to construct degenerate oligonucleotide primers allowing PCR amplification of cDNA or genomic DNA.

It has been shown that the cells from camel spleen imprint cells reacted with rabbit and anti camel immunoglobulin sera and that the spleen was hence a site of synthesis of at least one immunoglobulin class. cDNA was therefore synthetised from camel spleen mRNA. The conditions for the isolation of RNA were the following: total RNA was isolated from the dromedary spleen by the guanidium isothiocyanate method (15). mRNA was purified with oligo T-paramagnetic beads.

cDNA synthesis is obtained using 1 μg mRNA template, an oligodT primer and reverse transcriptase (BOERHINGER MAN). Second strand CDNA is obtained using RNAse H and *E. coli* DNA polymerase I according to the condition given by the supplier.

Relevant sequences were amplified by PCR: 5 ng of cDNA was amplified by PCR in a 100 μl reaction mixture (10 mM Tris-HCl pH 8.3, 50 mM KCl, 15 mM $MgCl_2$, 0.01% (w/v) gelatine, 200 μM of each dNTP and 25 pmoles of each primer) overlaid with mineral oil (Sigma).

Degenerate primers containing EcoRI and KpnI sites and further cloned into pUC 18. After a round of denaturing and annealing (94° C. for 5 min and 54° C. for 5 min), 2 units of Taq DNA polymerase were added to the reaction mixture before subjecting it to 35 cycles of amplification: 1 min at 94° C. (denature) 1 min at 54° C. (anneal), 2 min at 72° C. (elongate). To amplify DNA sequences between V_{HH} and C_H2 domains, (#72 clones), the PCR was performed in the same conditions with the exception that the annealing temperature was increased to 60° C.

One clone examined (#56/36) had a sequence corresponding to the N-terminal part of a C_H2 domain identical to the sequence of the 28 Kd fragment. The availability of this sequence data allowed the construction of an exact 3' primer and the cloning of the region between the N-terminal end of the V_{HH} and the C_H2 domain.

5' primers corresponding to the mouse V_{HH} (16) and containing a XhoI restriction site were used in conjunction with the 3' primer in which a KpnI site had been inserted and the amplified sequences were cloned into pBluescript^R. Clone #56/36 which displayed two internal HaeIII sites was digested with this enzyme to produce a probe to identify PCR positive clones.

After amplification the PCR products were checked on a 1.2% (w/v) agarose gel. Cleaning up of the PCR products included a phenol-chloroform extractio followed by further purification by HPLC (GEN-PAC FAX column, Waters) and finally by using the MERMAID or GENECLEAN II kit, BIO 101, Inc) as appropriate. After these purification steps, the amplified cDNA was then digested with EcoRI and KpnI for series #56 clones and with XhoI and KpnI for series #72 clones. A final phenol-chloroform extraction preceded the ligation into pUC 18(series=56 clones) or into pBluescript^R (series=72 clones).

All the clones obtained were smaller that the 860 base pairs to be expected if they possessed a complet V_{HH} and C_H1 region. Partial sequence data corresponding to the N-terminal of the V_{HH} region reveals that out of 20 clones, 3 were identical and possibly not independent. The sequences obtained ressemble the human subgroup III and the murine subgroups IIIa and IIIb (Table 2).

Clones corresponding to two different sets of C_H2 protein sequences were obtained. A first set of sequences (=72/41) had a N-terminal C_H2 region identical to the one obtained by protein sequencing of the 28 Kd papain fragments of the γ3 heavy chain, a short hinge region containing 3 cysteines and a variable region corresponding to the framework (FR4) residues encoded by the J minigenes adjoining the hinge. The C_H1 domain is entirely lacking. This cDNA corresponds to the γ3 chain (Table 4).

In one closely related sequence (#72/1) the proline in position 259 is replaced by threonine.

The sequence corresponding to the C_H3 and the remaining part of the C_H2 was obtained by PCR of the CDNA using as KpnI primer a poly T in which a KpnI restriction site had been inserted at the 5' end. The total sequence of the γ3 chain corresponds to a molecular weight (MW) which is in good agreement with the data obtained from SDS PAGE electrophoresis.

The sequence of this γ3 chain presents similarities with other γ chains except that it lacks the C_H1 domain, the V_{HH} domain being adjacent to the hinge.

One or all three of the cysteines could be probably responsible for holding the two γ3 chains together.

These results have allowed us to define a model for the IgG_3 molecule based on sequence and papain cleavage (FIG. 5).

Papain can cleave the molecule on each side of the hinge disulfides and also between C_H2 and C_H3. Under non reducing conditions the V_{HH} domains of IgG_3 can be iso-

lated as disulfide linked dimer or as monomer depending on the site of papain cleavage.

A second set of clones #72/29 had a slightly different sequence for the C_H2 and was characterized by a very long hinge immediately preceded by the variable domain. This hinge region has 3 cysteines at its C-terminal end in a sequence homologeous to the γ3 hinge. Such second set of clones could represent the IgG_2 subclass. For the constant part of the γ3 and also for the putative γ2, most clones are identical showing the γ2 or γ3 specific sequences. A few clones such as #72/1 however show minor differences. For instance in the case of clones #72/1 two nucleotide differences are detected.

Several V_{HH} regions cDNA's have now been totally or partially sequenced with the exception of a short stretch at the N-terminal end which is primer derived.

Upon translation the majority shows the characteristic heavy chain Ser_{21} Cys_{22} and Tyr_{90} Tyr_{91} Cys_{92} sequences, of the intra V_{HH} region disulfide bridge linking residues 22 and 92. All these clones have a sequence corresponding to the framework 4 (FR4) residues of the variable region immediately preceding the postulated hinge sequence (Table 3). This sequence is generated by the J minigenes and is in the majority of cases similar to the sequence encoded by the human and murine J minigenes. The sequence length between region Cys_{92} and the C-terminal end of the V_{HH} regions is variable and, in the sequences determined, range from 25 to 37 amino-acids as one might expect from the rearrangements of J and D minigenes varying in length.

Several important questions are raised by the sole existence of these heavy chain immunoglobulins in a non pathological situation. First of all, are they bonafide antibodies? The heavy chain immunoglobulins obtained from trypanosome infected camels react with a large number of parasite antigens as shown in part I of these examples. This implies that the camelid immune system generates an extensive number of binding sites composed of single V_{HH} domains. This is confirmed by the diversity of the V_{HH} regions of the heavy chain immunogobulins obtained by PCR.

The second question is "how are they secreted?". The secretion of immunoglobulin heavy chains composing four-chain model immunoglobulins does not occur under normal conditions. A chaperoning protein, the heavy chain binding protein, or BIP protein, prevents heavy chains from being secreted. It is only when the light chain dispplaces the BIP protein in the endoplasmatic reticulum that secretion can occur (13).

The heavy chain dimer found in the serum of human or mice with the so-called "heavy chain disease" lack the C_H1 domains thought to harbour the BIP site (14). In the absence of thi domain the BIP protein can no longer bind and prevent the transport of the heavy chains.

The presence in camels of a IgG1 class composed of heavy and light chains making up between 25% and 50% of the total IgG molecules also raises the problem as to how maturation and class switching occurs and what the role of the light chain is. The camelid light chain appears unusually large and heterogeneous when examined in SDS PAGE.

The largest dimension of an isolated domain is 40 Å and the maximum attainable span between binding sites of a conventional IgG with C_H1 and V_{HH} will be of the order of 160 Å ($2V_{HH}+2C_H1$) (19). The deletion of C_H1 domain in the two types of heavy chain antibodies devoid of light chains, already sequenced has, as a result, a modification of this maximum span (FIG. 6). In the IgG3 the extreme distance between the extremities of the V_{HH} regions will be of the order of 80 Å ($2V_{HH}$). This could be a severe limitation for agglutinating or cross linking. In the IgG_2 this is compensated by the extremely long stretch of hinge, composed of a 12-fold repeat of the sequence Pro-X (where X is Gln, Lys or Glu) and located N-terminal to the hinge disulfide bridges. In contrast, in the human IgG3, the very long hinge which also apparently arose as the result of sequence duplication does not contribute to increase the distance spanning the two binding sites as this hinge is inter-spersed with disulfide bridges.

The single V_{HH} domain could also probably allow considerably rotational freedom of the binding site versus the Fc domain.

Unlike myeloma heavy chains which result probably from C_H1 deletion in a single antibody producing cell, or heavy chain antibodies produced by expression cloning (15); the camelid heavy chain antibodies (devoid of light chains) have emerged in a normal immunological environment and it is expected that they will have undergone the selective refinement in specificity and affinity accompanying B cell maturation.

Expression and purification of the camel $V_{HH}21$ (DR21 on FIG. 7) protein from *E. coli*

The clones can be expressed in several types of expression vectors. As an example using a commercially available vector Immuno PBS (Huse et al: Science (1989) 246, 1275), clones produced in Bluescript® according to the above described procedure, have been recovered by PCR using the same XhoI containing $_5$' primer and a new 3' primer, corresponding to residues 113–103 in the framework of the immunoglobulins, in which an Spe site has been constructed: TC TTA ACT AGT GAG GAG ACG GTG ACC TG SEQ ID NO: 51. This procedure allowed the cloning of the V_{HH} in the Xho/Spe site of the Immuno PBS vector. However, the 3' end of the gene was not in phase with the identification "tag" and the stop codon of the vector. To achieve this, the construct was cut with Spe and the 4 base overhangs were filled in, using the Klenow fragment after which the vector was religated.

The expression vector plasmid ipBS (immunopBS) (Stratacyte) contains a pel B leader sequence which is used for immunoglobulin chain expression in *E. coli* under the promotor pLAC control, a ribosome binding site, and stop codons. In addition, it contains a sequence for a c-terminal decapeptide tag.

E. coli JM101 harboring the ipBS-$V_{HH}21$ plasmid was grown in 1 l of TB medium with 100 μg/ml ampicillin and 0.1% glucose at 32° C. Expression was induced by the addition of 1 mM IPTG (final concentration) at an OD_{550} of 1.0. After overnight induction at 28° C., the cells were harvested by centrifugation at 4.000 g for 10 min (4° C.) and resuspended in 10 ml TES buffer (0.2M Tris-HCL pH 8.0, 0.5 mM EDTA, 0.5M sucrose). The suspension was kept on ice for 2 hours. Periplasmic proteins were removed by osmotic shock by addition of 20 ml TES buffer diluted 1:4 v/v with water, kept on ice for one hour and subsequently centrifugated at 12.000 g for 30 min. at 4° C. The supernatant periplasmic fraction was dialysed against Tris-HCl pH 8.8, NaCl 50 mM, applied on a fast Q Sepharose flow (Pharmacia) column, washed with the above buffer prior and eluted with a linear gradient of 50 mM to 1M NaCl in buffer.

Fractions containing the V_{HH} protein were further purified on a Superdex 75 column (Pharmacia) equilibrated with PBS buffer (0.01M phosphate pH 7.2, 0.15M NaCl). The yield of purified VHI protein varies from 2 to 5 mg/l cell culture.

Fractions were analyzed by SDS-PAGE(I). Positive identification of the camel V_{HH} antibody fragment was done by

Western Blot analysis using antibody raised in rabbits against purified camel $IgGH_3$ and an anti-rabbit IgG-alkaline phosphatase conjugate (II).

As protein standards (Pharmacia) periplasmic proteins prepared from 1 ml of IPTG-induced JM101/ipBS $V_{HH}21$ were used. FIG. 8 shows: C.D:fractions from fast S Sepharose column chromatography (C:Eluted at 650 mM NaCl D:Eluted at 700 mM NaCl) E.F:fractions from Superdex 75 column chromatography.

As can be seen, the major impurity is eliminated by ionexchange chromatography and the bulk of the remaining impurities are eliminated by gel filtration.

TABLE 1

Comparison of the N terminal Camel C_H2 and C_H3 sequences with the translated cDNA sequences of Camel immunoglobulins and with the corresponding human γ sequences. (Numbering according to Kabat et al (1987)(7).

		250										260										270						
Camel	γ_3 28Kd	–	L	P	G	G	P	S	V	F	V	F	P	P	K	P	K	D	V	L	S	I	X	G	X	P	SEQ ID NO:54	– –
Clone	#72/1	–	L	P	G	G	P	S	V	F	V	F	P	T	K	P	K	D	V	L	S	I	S	G	R	P	SEQ ID NO:55	– –
Clone	#72/4	–	L	P	G	G	P	S	V	F	V	F	P	P	K	P	K	D	V	L	S	I	S	G	R	P	SEQ ID NO:56	– –
Clone	#72/29	–	L	L	G	G	P	S	V	F	I	F	P	P	K	P	K	D	V	L	S	I	S	G	R	P	SEQ ID NO:57	– –
Human	$\gamma_1\gamma_3$	–	L	L	G	G	P	S	V	F	L	F	P	P	K	P	K	D	T	L	M	I	S	R	T	P	SEQ ID NO:112	– –
C_H2	γ_2	–	V	A	–	G	P	S	V	F	L	F	P	P	K	P	K	D	T	L	M	I	S	R	T	P	SEQ ID NO:113	– –
	γ_4	–	F	L	G	G	P	S	V	F	L	F	P	P	K	P	K	D	T	L	M	I	S	R	T	P	SEQ ID NO:114	– –

		C_H2 | C_H3																										
		360									370																	
Camel	γ_3 14Kd	–	K	G	Q	T	R	E	P	Q	V	Y	T	L	A	P	X	R	L	E	L	SEQ ID NO:54	– –
Human	γ_1	–	K	G	Q	P	R	E	P	Q	V	Y	T	L	P	P	S	R	D	E	L	SEQ ID NO:115	– –
C_H2/C_H3	γ_2, γ_3	–	K	G	Q	P	R	E	P	Q	V	Y	T	L	P	P	S	R	E	E	M	SEQ ID NO:116	– –
	γ_4	–	K	G	Q	P	R	E	P	Q	V	Y	T	L	P	P	S	Q	E	E	M	SEQ ID NO:117	– –

TABLE 2

A comparison of N Terminal Fr 1 regions of Camel V_{HH} with a Human V_HIII subgroup protein and a mouse V_HIIIA subgroup protein. The invariable subgroup specific residues are grayed.

	10									20									30									
Primer Derived	G	G	S	V	Q	T	G	G	S	L	R	L	S	C	E	I	S	G	L	T	F	D	#72/4	SEQ ID NO:1
	G	G	S	V	Q	T	G	G	S	L	R	L	S	C	A	V	S	G	F	S	F	S	#72/3	SEQ ID NO:2
	G	G	S	E	Q	G	G	G	S	L	R	L	S	C	A	I	S	G	Y	T	Y	G	#72/7	SEQ ID NO:3
	G	G	S	V	Q	P	G	G	S	L	T	L	S	C	T	V	S	G	A	T	Y	S	#72/17	SEQ ID NO:4
	G	G	S	V	Q	A	G	G	S	L	R	L	S	C	T	G	S	G	F	P	Y	S	#72/18	SEQ ID NO:5
D	V	Q	L	V	A	S	G	G	G	S	V	G	A	G	G	S	L	R	L	S	C	T	A	S	G	D	S	F	S	#72/2	SEQ ID NO:58
E	V	K	L	V	E	S	G	G	G	L	V	E	P	G	G	S	L	R	L	S	C	A	T	S	G	F	T	F	S	Mouse V_HIII_A	SEQ ID NO:118
	E	V	Q	L	L	S	G	G	G	L	V	Q	P	G	G	S	L	R	L	S	C	A	A	S	G	F	T	F	S	Human V_HIII	SEQ ID NO:119

TABLE 3

Comparison of some Framework 4 residues found in the Camel V_{HH} region with the Framework 4 residues corresponding to the consensus region of the Human and Mouse J minigenes.

	FrameWork 4											J Genes
Human	W	G	Q	G	T	L	V	T	V	S	S	SEQ ID NO:9	J1, J4, J5
	W	G	R	G	T	L	V	T	V	S	S	SEQ ID NO:130	J2
	W	G	Q	G	T	T	V	T	V	S	S	SEQ ID NO:120	J6
	W	G	Q	G	T	M	V	T	V	S	S	SEQ ID NO:121	J3

TABLE 3-continued

Comparison of some Framework 4 residues found in the Camel V_{HH} region with the Framework 4 residues corresponding to the consensus region of the Human and Mouse J minigenes.

	Residues	SEQ ID	
Murine	W G Q G T T L T V S S	SEQ ID NO:122	J1
	W G Q G T L V T V S S	SEQ ID NO:9	J2
	W G Q G T S V T V S S	SEQ ID NO:123	J3
	W G A G T T V T V S S	SEQ ID NO:124	J4
			cDNA Clones
Camel	W G Q G T Q V T V S S	SEQ ID NO:8	Clones
	W G Q G T Q V T V S S	SEQ ID NO:8	#72/19 = #72/3
	W G Q G T L V T V S S	SEQ ID NO:9	1 Clone
	W G R G T Q V T V S S	SEQ ID NO:59	#72/24
	W G Q G T H V T V S S	SEQ ID NO:60	#72/21
	W G Q G I Q V T A S S	SEQ ID NO:61	#72/16

TABLE 4

	CDR3		SEQ ID	
	d y y g s s - - - - - - - - y - - f - - - - - - d v 95 100 a b c d e f g h i j k - - - - 101	W G A G T T V T V S S 105 110	SEQ ID NO:12;	Mouse $V_{H_{III}}$ sequence
1	a l q p g g y c g y g x - - - - - - - - - - - - - c l	W G Q G T Q V T V S S	SEQ ID NO:13	
2	v s l m d r i s q h - - - - - - - - - - - - - - - g c	R G Q G T Q V T V S L	SEQ ID NO:14	
3	v p a h l g p g n i l d l k k y - - - - - - - - - k y	W G Q G T Q V T V S S	SEQ ID NO:15	
4	f c y s t a g d g g s g e - - - - - - - - - - - - m y	W G Q G T Q V T V S S	SEQ ID NO:16	
7	e l s g g s c e l p l l f - - - - - - - - - - - - d y	W G Q G T Q V T V S S	SEQ ID NO:17	
9	d w k y w t c g a q t g g y f - - - - - - - - - - g q	W G Q G A Q V T V S S	SEQ ID NO:18	
11	r l t e m g a c d a r w a t l a t r t f a y n y	W G Q G T Q V T V S S	SEQ ID NO:19	Random sample
13	q k k d r t r w a e p r e w - - - - - - - - - - - n n	W G Q G T Q V T A S S	SEQ ID NO:20	
16	g s r f s s p v g s t s r l e s - s d y - - n y	W G Q G I Q V T A S S	SEQ ID NO:21	
17	a d p s i y y s i l x i e y - - - - - - - - - - - k y	W G Q G T Q V T V S S	SEQ ID NO:22	18 different camel
18	d s p c y m p t m p a p p i r d s f g w - - d d	F G Q G T O V T V S S	SEQ ID NO:23	VH_H region
19	t s s t y w y c t t a p y - - - - - - - - - - - - n v	W G Q G T Q V T V S S	SEQ ID NO:24	
20	t e i e w y g c n l r t t f - - - - - - - - - - - t r	W G Q G T Q V T V S S	SEQ ID NO:25	
21	n q l a g g w y l d p n y w l s v g a y - - a i	W G Q G T H V T V S S	SEQ ID NO:26	
24	r l t e m g a c d a r w a t l a t r t f a y n y	W G R G T Q V T V S S	SEQ ID NO:27	
25	d g w t r k e g g i g l p w s v q c e d g y n y	W G Q G T Q V T V S S	SEQ ID NO:28	
27	d s y p c h l l - - - - - - - - - - - - - - - - d v	W G Q G T Q V T V S S	SEQ ID NO:29	
29	v e y p i a d m c s - - - - - - - - - - - - - - - r y	G D P G T Q V T V S S	SEQ ID NO:30	
	CDR3			

Human & mouse - size range	0–19 ea	over 600 entries
Camel	8–24 ea	18 entries

TABLE 5

EVQLVESGGG LVQP	GGSLRL SCAASG	SEQ ID NO:80	CDR1	WVRQA PGKGLEWVS	SEQ ID NO:81	CDR2
GG SVQGGGSLRL SCA I SG	SEQ ID NO:84	CDR1	WE REG PGKE REG I A	SEQ ID NO:85	CDR2
GG SVQAGGSLRL SCAS SS	SEQ ID NO:88	CDR1	WYRQA PGKE REF VS	SEQ ID NO:89	CDR2

RFTIS RDNSKNTLYL QMNSLRAEDTAVY YCAR	SEQ ID NO:82	CDR3	WGQGTLVT VSS	SEQ ID NO:83
RFTIS QDSTLKTMYL LMNNLKPEDTGTY YCAA	SEQ ID NO:86	CDR3	WGQGTQVT VSS	SEQ ID NO:91
RFTIS QDSAKNTVYL QMNSLKPEDTAMY YCKI	SEQ ID NO:90	CDR3	WGQGTQVT VSS	SEQ ID NO:87

	camel V_{HH}	hinge	C_H2	
camel	WGQGTQVT VSS	GTNEVCKCPKCP	APELPGG PSVFVFP	SEQ ID NO:91
	WGQGTQVT VSS	EPKIPQPQPKPQPQP		
		QPQPKPQP		
		KPEPECTCPKCP	APELLGG PSVFIFP	SEQ ID NO:87

	human C_H1	hinge	C_H2	
human gamma 3	KVDKRV	ELKTPLGDTTHTCPRCP		
		EPKCSDTPPPCPRCP		
		EPKSCDTPPPCPRCP	APELLGG PSVFLFP	SEQ ID NO:126
human gamma 1	KVDKK	AEPKSCDKTHTCPPCP	APELLGG PSVFLFP	SEQ ID NO:127
human gamma 2	KVKVTV	ERKCCVECPPCP	APPVAG - PSVFLFP	SEQ ID NO:128
human gamma 4	KVDKRV	ESKYGPPCPSCP	APEFLGG PSVFLFP	SEQ ID NO:129

REFERENCES

1. Ward, E. S., Gussow, D., Griffits, A. D., Jones, P. T. and Winter G. Nature 341, 544–546 (1989).
2. Ungar-Waron H., Eliase E., Gluckman A. and Trainiii Z. Isr. J. Vet. Med., 43, 198–203 (1987).
3. Bajyana Songa E. and Hamers R., Ann. Soc. Belge Med. trop., 68, 233–240 (1988).
4. Edelman G. M., Olins D. E., Gally J. A. and Zinder N. D., Proc. Nat. Acad. Sc., 50, 753 (1963).
5. Franek F. and Nezlin R. S., Biokhimiya, 28, 193 (1963).
6. Roitt I. M., Brostof J. and Male D. K., Immunology, Gower Med. Pub. London. New-York, p.9.2. (1985).
7. Schiffer M., Girling R. L., Ely K. R. and Edmundson B., Biochemistry, 12, 4620–4631 (1973).
8. Fleischman J. B., Pain R. H. and Porter R. R., Arch. Biochem. Biophys, Suppl. 1, 174 (1962).
9. Roholt O., Onoue K. and Pressman D., PNAS 51, 173–178 (1964).
10. Seligmann M., Mihaesco E., Preud'homme J. L., Danon F. and Brouet J. C., Immunological Rev., 48, 145–167 (1979).
11. Henderschot L., Bole D., Kohler G. and Kearney J. F., The Journal of Cell Biology, 104, 761–767 (1987).
12. Henderschot L. M., The Journal of Cell Biology, 111, 829–837 (1990).
13. Hamers-Casterman, C., E. Wittouck, W. Van der Leo and R. Hamers, Journal of Inmunogenetics, 6, 373–381 (1979).
14. Applied Biosystems—Ethanol Precipitation of Electro Eluted Electrodialysed Sample. Issue n° 27.
15. Maniatis, T. E. F. Fritsch and J. Sambrook, Molecular Cloning. A Laboratory Manual (1988).
16. Sastry et al., PNAS, 86, 5728, (1989).
17. Sanger, F., S. Nicklen and A. R. Coulson, Proc. Natl. Acad. Sci., U.S.A., 74, 5463–5467 (1977).
18. Kabat E. A., Tai Te Wu, M. Reid-Miller, H. M. Perry and K. S. Gottesman, U.S. Dpt of Health and Human Services, Public Health Service, National Institutes of Health (1987).
19. Valentine, R. C. and N. M. Geen, J. M. B., 27, 615–617 (1967).

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i i i) NUMBER OF SEQUENCES: 130

-continued

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..22
(D) OTHER INFORMATION: /label=FRAMEWORK 1

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Gly Gly Ser Val Gln Thr Gly Gly Ser Leu Arg Leu Ser Cys Glu Ile
1 5 10 15

Ser Gly Leu Thr Phe Asp
20

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..22
(D) OTHER INFORMATION: /label=FRAMEWORK 1

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Gly Gly Ser Val Gln Thr Gly Gly Ser Leu Arg Leu Ser Cys Ala Val
1 5 10 15

Ser Gly Phe Ser Phe Ser
20

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..22
(D) OTHER INFORMATION: /label=FRAMEWORK 1

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Gly Gly Ser Glu Gln Gly Gly Gly Ser Leu Arg Leu Ser Cys Ala Ile
1 5 10 15

Ser Gly Tyr Thr Tyr Gly
20

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

-continued

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..22
(D) OTHER INFORMATION: /label=FRAMEWORK 1

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Gly Gly Ser Val Gln Pro Gly Gly Ser Leu Thr Leu Ser Cys Thr Val
1               5                  10                  15

Ser Gly Ala Thr Tyr Ser
            20
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..22
(D) OTHER INFORMATION: /label=FRAMEWORK 1

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
Gly Gly Ser Val Gln Ala Gly Gly Ser Leu Arg Leu Ser Cys Thr Gly
1               5                  10                  15

Ser Gly Phe Pro Tyr Ser
            20
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..21
(D) OTHER INFORMATION: /label=FRAMEWORK 1

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
Gly Gly Ser Val Gln Ala Gly Gly Ser Leu Arg Leu Ser Cys Val Ala
1               5                  10                  15

Gly Phe Gly Thr Ser
            20
```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..21
(D) OTHER INFORMATION: /label=FRAMEWORK 1

-continued

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Gly Gly Ser Val Gln Ala Gly Gly Ser Leu Arg Leu Ser Cys Val Ser
1 5 10 15

Phe Ser Pro Ser Ser
20

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..11
(D) OTHER INFORMATION: /label=FRAMEWORK 4

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Trp Gly Gln Gly Thr Gln Val Thr Val Ser Ser
1 5 10

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..11
(D) OTHER INFORMATION: /label=FRAMEWORK 4

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
1 5 10

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..11
(D) OTHER INFORMATION: /label=FRAMEWORK 4

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Trp Gly Gln Gly Ala Gln Val Thr Val Ser Ser
1 5 10

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

-continued

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..11
(D) OTHER INFORMATION: /label=FRAMEWORK 4

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```
Trp Gly Gln Gly Thr Gln Val Thr Ala Ser Ser
1               5                  10
```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..11
(D) OTHER INFORMATION: /label=FRAMEWORK 4

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
Arg Gly Gln Gly Thr Gln Val Thr Val Ser Leu
1               5                  10
```

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..14
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```
Ala Leu Gln Pro Gly Gly Tyr Cys Gly Tyr Gly Xaa Cys Leu Trp Gly
1               5                  10                  15

Gln Gly Thr Gln Val Thr Val Ser Ser
            20                  25
```

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..12
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```
Val Ser Leu Met Asp Arg Ile Ser Gln His Gly Cys Arg Gly Gln Gly
1               5                  10                  15

Thr Gln Val Thr Val Ser Leu
            20
```

-continued

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..18
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```
Val Pro Ala His Leu Gly Pro Gly Ala Ile Leu Asp Leu Lys Lys Tyr
1               5                  10                  15

Lys Tyr Trp Gly Gln Gly Thr Gln Val Thr Val Ser Ser
            20                  25
```

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..15
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```
Phe Cys Tyr Ser Thr Ala Gly Asp Gly Gly Ser Gly Glu Met Tyr Trp
1               5                  10                  15

Gly Gln Gly Thr Gln Val Thr Val Ser Ser
            20                  25
```

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..15
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```
Glu Leu Ser Gly Gly Ser Cys Glu Leu Pro Leu Leu Phe Asp Tyr Trp
1               5                  10                  15

Gly Gln Gly Thr Gln Val Thr Val Ser Ser
            20                  25
```

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single

-continued

(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..17
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```
Asp Trp Lys Tyr Trp Thr Cys Gly Ala Gln Thr Gly Gly Tyr Phe Gly
1               5                  10                  15

Gln Trp Gly Gln Gly Ala Gln Val Thr Val Ser Ser
            20                  25
```

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..24
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```
Arg Leu Thr Glu Met Gly Ala Cys Asp Ala Arg Trp Ala Thr Leu Ala
1               5                  10                  15

Thr Arg Thr Phe Ala Tyr Asn Tyr Trp Gly Gln Gly Thr Gln Val Thr
            20                  25                  30

Val Ser Ser
        35
```

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..16
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```
Glu Lys Lys Asp Arg Thr Arg Trp Ala Glu Pro Arg Glu Trp Asn Asn
1               5                  10                  15

Trp Gly Gln Gly Thr Gln Val Thr Ala Ser Ser
            20                  25
```

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:

-continued

(A) NAME/KEY: Domain
(B) LOCATION: 1..21
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Gly Ser Arg Phe Ser Ser Pro Val Gly Ser Thr Ser Arg Leu Glu Ser
1 5 10 15

Ser Asp Tyr Asn Tyr Trp Gly Gln Gly Thr Gln Val Thr Ala Ser Ser
20 25 30

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..16
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Ala Asp Pro Ser Ile Tyr Tyr Ser Ile Leu Xaa Ile Glu Tyr Lys Tyr
1 5 10 15

Trp Gly Gln Gly Thr Gln Val Thr Val Ser Ser
20 25

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..22
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Asp Ser Pro Cys Tyr Met Pro Thr Met Pro Ala Pro Pro Ile Arg Asp
1 5 10 15

Ser Phe Gly Trp Asp Asp Phe Gly Gln Gly Thr Gln Val Thr Val Ser
20 25 30

Ser

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..15
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:24:

-continued

```
Thr Ser Ser Phe Tyr Trp Tyr Cys Thr Thr Ala Pro Tyr Asn Val Trp
1               5                  10                  15

Gly Gln Gly Thr Gln Val Thr Val Ser Ser
            20                  25
```

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..16
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:25:

```
Thr Glu Ile Glu Trp Tyr Gly Cys Asn Leu Arg Thr Thr Phe Thr Arg
1               5                  10                  15

Trp Gly Gln Gly Thr Gln Val Thr Val Ser Ser
            20                  25
```

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..22
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:26:

```
Asn Gln Leu Ala Gly Gly Trp Tyr Leu Asp Pro Asn Tyr Trp Leu Ser
1               5                  10                  15

Val Gly Ala Tyr Ala Ile Trp Gly Gln Gly Thr His Val Thr Val Ser
            20                  25                  30

Ser
```

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..24
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:27:

```
Arg Leu Thr Glu Met Gly Ala Cys Asp Ala Arg Trp Ala Thr Leu Ala
1               5                  10                  15

Thr Arg Thr Phe Ala Tyr Asn Tyr Trp Gly Arg Gly Thr Gln Val Thr
            20                  25                  30
```

-continued

Val Ser Ser
35

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..24
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Asp Gly Trp Thr Arg Lys Glu Gly Gly Ile Gly Leu Pro Trp Ser Val
1 5 10 15

Gln Cys Glu Asp Gly Tyr Asn Tyr Trp Gly Gln Gly Thr Gln Val Thr
20 25 30

Val Ser Ser
35

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..10
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Asp Ser Tyr Pro Cys His Leu Leu Asp Val Trp Gly Gln Gly Thr Gln
1 5 10 15

Val Thr Val Ser Ser
20

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..12
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Val Glu Tyr Pro Ile Ala Asp Met Cys Ser Arg Tyr Gly Asp Pro Gly
1 5 10 15

Thr Gln Val Thr Val Ser Ser
20

(2) INFORMATION FOR SEQ ID NO:31:

-continued

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..27
(D) OTHER INFORMATION: /label=CH2

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:31:

```
Ala Pro Glu Leu Leu Gly Gly Pro Thr Val Phe Ile Phe Pro Pro Lys
1               5                  10                  15

Pro Lys Asp Val Leu Ser Ile Thr Leu Thr Pro
            20                  25
```

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..27
(D) OTHER INFORMATION: /label=CH2

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:32:

```
Ala Pro Glu Leu Pro Gly Gly Pro Ser Val Phe Val Phe Pro Thr Lys
1               5                  10                  15

Pro Lys Asp Val Leu Ser Ile Ser Gly Arg Pro
            20                  25
```

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..27
(D) OTHER INFORMATION: /label=CH2

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:33:

```
Ala Pro Glu Leu Pro Gly Gly Pro Ser Val Phe Val Phe Pro Pro Lys
1               5                  10                  15

Pro Lys Asp Val Leu Ser Ile Ser Gly Arg Pro
            20                  25
```

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

-continued

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..27
(D) OTHER INFORMATION: /label=CH2

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Ile Phe Pro Pro Lys
1 5 10 15

Pro Lys Asp Val Leu Ser Ile Ser Gly Arg Pro
20 25

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..12
(D) OTHER INFORMATION: /label=CH3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Gly Gln Thr Arg Glu Pro Gln Val Tyr Thr Leu Ala
1 5 10

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..18
(D) OTHER INFORMATION: /label=CH3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Gly Gln Thr Arg Glu Pro Gln Val Tyr Thr Leu Ala Pro Xaa Arg Leu
1 5 10 15

Glu Leu

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Region
(B) LOCATION: 1..12
(D) OTHER INFORMATION: /label=hinge

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Gly Thr Asn Glu Val Cys Lys Cys Pro Lys Cys Pro
1 5 10

-continued

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Region
(B) LOCATION: 1..35
(D) OTHER INFORMATION: /label=hinge

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Glu Pro Lys Ile Pro Gln Pro Gln Pro Lys Pro Gln Pro Gln Pro Gln
1 5 10 15

Pro Gln Pro Lys Pro Gln Pro Lys Pro Glu Pro Glu Cys Thr Cys Pro
20 25 30

Lys Cys Pro
35

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..28
(D) OTHER INFORMATION: /label=CH2

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:39:

Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Val Phe Pro Pro Lys
1 5 10 15

Pro Lys Asp Val Leu Ser Ile Ser Gly Xaa Pro Lys
20 25

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..28
(D) OTHER INFORMATION: /label=CH2

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Ala Pro Glu Leu Pro Gly Gly Pro Ser Val Phe Val Phe Pro Thr Lys
1 5 10 15

Pro Lys Asp Val Leu Ser Ile Ser Gly Arg Pro Lys
20 25

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 amino acids
(B) TYPE: amino acid

-continued

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..28
(D) OTHER INFORMATION: /label=CH2

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Ala Pro Glu Leu Pro Gly Gly Pro Ser Val Phe Val Phe Pro Pro Lys
1 5 10 15

Pro Lys Asp Val Leu Ser Ile Ser Gly Arg Pro Lys
20 25

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..28
(D) OTHER INFORMATION: /label=CH2

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:42:

Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Ile Phe Pro Pro Lys
1 5 10 15

Pro Lys Asp Val Leu Ser Ile Ser Gly Arg Pro Lys
20 25

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:43:

Val Thr Val Ser Ser Gly Thr Asn Glu Val Cys Lys Cys Pro Lys Cys
1 5 10 15

Pro Ala Pro Glu Leu Pro Gly Gly Pro Ser Val Phe Val Phe Pro
20 25 30

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 54 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Region
(B) LOCATION: 1..54
(D) OTHER INFORMATION: /label=hinge

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:44:

Val Thr Val Ser Ser Glu Pro Lys Ile Pro Gln Pro Gln Pro Lys Pro

-continued

```
1               5                  10                  15

Gln Pro Gln Pro Gln Pro Gln Pro Lys Pro Gln Pro Lys Pro Glu Pro
            20                  25                  30

Glu Cys Thr Cys Pro Lys Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro
        35                  40                  45

Ser Val Phe Ile Phe Pro
    50
```

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Region
(B) LOCATION: 1..14
(D) OTHER INFORMATION: /label=hinge

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..14
(D) OTHER INFORMATION: /label=CH2

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:45:

```
Ala Pro Glu Leu Pro Gly Gly Pro Ser Val Phe Val Phe Pro
1               5                  10
```

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..14
(D) OTHER INFORMATION: /label=CH2

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:46:

```
Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Ile Phe Pro
1               5                  10
```

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: Other
(A) DESCRIPTION: DNA (synthetic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:47:

```
CGCCATCAAG GTAACAGTTG A                                    21
```

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid

-continued

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: Other
(A) DESCRIPTION: DNA (synthetic)

(i x) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 12..17
(D) OTHER INFORMATION: /label=XhoI site
/ note= "RESTRICTION SITE"

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:48:

AGGTCCAGCT GCTCGAGTCT GG 22

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: Other
(A) DESCRIPTION: DNA (synthetic)

(i x) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 12..17
(D) OTHER INFORMATION: /label=XhoI site
/ note= "Restriction site"

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:49:

AGCTCCAGCT GCTCGAGTCT GG 22

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: Other
(A) DESCRIPTION: DNA (synthetic)

(i x) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 12..17
(D) OTHER INFORMATION: /label=XhoI site
/ note= "restriction site"

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:50:

AGGTCCAGCT TCTCGAGTCT GG 22

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: Other
(A) DESCRIPTION: DNA (synthetic)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:51:

TCTTAACTAG TGAGGAGACG GTGACCTG 28

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:

-continued

(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: Other
(A) DESCRIPTION: DNA (synthetic)

(i x) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1..5
(D) OTHER INFORMATION: /label=SpeI

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:52:

CTAGTGCACC ACCATCACCA TCACTAATAG 30

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: Other
(A) DESCRIPTION: DNA (synthetic)

(i x) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1..30
(D) OTHER INFORMATION: /note= "Sequence complementary to SEQ ID NO: 52"

(i x) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 26..30
(D) OTHER INFORMATION: /label=EcoRI

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:53:

ACGTGGTGGT AGTGGTAGTG ATTATCTTAA 30

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

(v i) ORIGINAL SOURCE:
(A) ORGANISM: Camelus dromedarius

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..25
(D) OTHER INFORMATION: /label=CH2

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 26..43
(D) OTHER INFORMATION: /label=CH3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:54:

Leu Pro Gly Gly Pro Ser Val Phe Val Phe Pro Pro Lys Pro Lys Asp
1 5 10 15

Val Leu Ser Ile Xaa Gly Xaa Pro Lys Gly Gln Thr Arg Glu Pro Gln
20 25 30

Val Tyr Thr Leu Ala Pro Xaa Arg Leu Glu Leu
35 40

-continued

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

(v i) ORIGINAL SOURCE:
(A) ORGANISM: Camelus dromedarius

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..24
(D) OTHER INFORMATION: /label=CH2
/ note= "Clone #72/1"

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:55:

Leu Pro Gly Gly Pro Ser Val Phe Val Phe Pro Thr Lys Pro Lys Asp
1 5 10 15

Val Leu Ser Ile Ser Gly Arg Pro
20

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

(v i) ORIGINAL SOURCE:
(A) ORGANISM: Camelus dromedarius

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..24
(D) OTHER INFORMATION: /label=CH2

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:56:

Leu Pro Gly Gly Pro Ser Val Phe Val Phe Pro Pro Lys Pro Lys Asp
1 5 10 15

Val Leu Ser Ile Ser Gly Arg Pro
20

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

(v i) ORIGINAL SOURCE:
(A) ORGANISM: Camelus dromedarius

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..24
(D) OTHER INFORMATION: /label=CH2

-continued

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:57:

```
Leu Leu Gly Gly Pro Ser Val Phe Ile Phe Pro Pro Lys Pro Lys Asp
1               5                  10                  15

Val Leu Ser Ile Ser Gly Arg Pro
            20
```

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

(v i) ORIGINAL SOURCE:
(A) ORGANISM: Camelus dromedarius

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..30
(D) OTHER INFORMATION: /label=Framework 1
/ note= "CAMEL"

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:58:

```
Asp Val Gln Leu Val Ala Ser Gly Gly Gly Ser Val Gly Ala Gly Gly
1               5                  10                  15

Ser Leu Arg Leu Ser Cys Thr Ala Ser Gly Asp Ser Phe Ser
            20                  25                  30
```

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(v i) ORIGINAL SOURCE:
(A) ORGANISM: Camelus dromedarius

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..11
(D) OTHER INFORMATION: /label=Framework 4

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:59:

```
Trp Gly Arg Gly Thr Gln Val Thr Val Ser Ser
1               5                  10
```

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(v i) ORIGINAL SOURCE:
(A) ORGANISM: Camelus dromedarius

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..11

-continued

(D) OTHER INFORMATION: /label=Framework 4

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:60:

Trp Gly Gln Gly Thr His Val Thr Val Ser Ser
1 5 10

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..11
(D) OTHER INFORMATION: /label=Framework 4

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:61:

Trp Gly Gln Gly Ile Gln Val Thr Ala Ser Ser
1 5 10

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(i x) FEATURE:
(A) NAME/KEY: Region
(B) LOCATION: 1..14
(D) OTHER INFORMATION: /label=VH

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..14
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:62:

Ala Leu Gln Pro Gly Gly Tyr Cys Gly Tyr Gly Xaa Cys Leu
1 5 10

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(v i) ORIGINAL SOURCE:
(A) ORGANISM: Camelus dromedarius

(i x) FEATURE:
(A) NAME/KEY: Region
(B) LOCATION: 1..12
(D) OTHER INFORMATION: /label=VH

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..12
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:63:

-continued

```
Val Ser Leu Met Asp Arg Ile Ser Gln His Gly Cys
1               5                  10
```

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(v i) ORIGINAL SOURCE:
(A) ORGANISM: Camelus dromedarius

(i x) FEATURE:
(A) NAME/KEY: Region
(B) LOCATION: 1..18
(D) OTHER INFORMATION: /label=VH

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..18
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:64:

```
Val Pro Ala His Leu Gly Pro Gly Ala Ile Leu Asp Leu Lys Lys Tyr
1               5                  10                  15

Lys Tyr
```

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(v i) ORIGINAL SOURCE:
(A) ORGANISM: Camelus bactrianus

(i x) FEATURE:
(A) NAME/KEY: Region
(B) LOCATION: 1..15
(D) OTHER INFORMATION: /label=VH

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..15
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:65:

```
Phe Cys Tyr Ser Thr Ala Gly Asp Gly Gly Ser Gly Glu Met Tyr
1               5                  10                  15
```

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(v i) ORIGINAL SOURCE:
(A) ORGANISM: Camelus dromedarius

(i x) FEATURE:
(A) NAME/KEY: Region
(B) LOCATION: 1..15

-continued

(D) OTHER INFORMATION: /label=VH

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..15
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:66:

Glu Leu Ser Gly Gly Ser Cys Glu Leu Pro Leu Leu Phe Asp Tyr
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(i x) FEATURE:
(A) NAME/KEY: Region
(B) LOCATION: 1..17
(D) OTHER INFORMATION: /label=VH

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..17
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:67:

Asp Trp Lys Tyr Trp Thr Cys Gly Ala Gln Thr Gly Gly Tyr Phe Gly
1 5 10 15

Gln

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(i x) FEATURE:
(A) NAME/KEY: Region
(B) LOCATION: 1..24
(D) OTHER INFORMATION: /label=VH

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..24
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:68:

Arg Leu Thr Glu Met Gly Ala Cys Asp Ala Arg Trp Ala Thr Leu Ala
1 5 10 15

Thr Arg Thr Phe Ala Tyr Asn Tyr
20

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

-continued

(i x) FEATURE:
(A) NAME/KEY: Region
(B) LOCATION: 1..16
(D) OTHER INFORMATION: /label=VH

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..16
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:69:

```
Gln Lys Lys Asp Arg Thr Arg Trp Ala Glu Pro Arg Glu Trp Asn Asn
1               5                  10                  15
```

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(i x) FEATURE:
(A) NAME/KEY: Region
(B) LOCATION: 1..21
(D) OTHER INFORMATION: /label=VH

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..21
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:70:

```
Gly Ser Arg Phe Ser Ser Pro Val Gly Ser Thr Ser Arg Leu Glu Ser
1               5                  10                  15

Ser Asp Tyr Asn Tyr
            20
```

(2) INFORMATION FOR SEQ ID NO:71:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(i x) FEATURE:
(A) NAME/KEY: Region
(B) LOCATION: 1..16
(D) OTHER INFORMATION: /label=VH

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..16
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:71:

```
Ala Asp Pro Ser Ile Tyr Tyr Ser Ile Leu Xaa Ile Glu Tyr Lys Tyr
1               5                  10                  15
```

(2) INFORMATION FOR SEQ ID NO:72:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

-continued

(i x) FEATURE:
(A) NAME/KEY: Region
(B) LOCATION: 1..22
(D) OTHER INFORMATION: /label=VH

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..22
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:72:

```
Asp Ser Pro Cys Tyr Met Pro Thr Met Pro Ala Pro Pro Ile Arg Asp
1               5                  10                  15

Ser Phe Gly Trp Asp Asp
            20
```

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(i x) FEATURE:
(A) NAME/KEY: Region
(B) LOCATION: 1..15
(D) OTHER INFORMATION: /label=VH

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..15
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:73:

```
Thr Ser Ser Phe Tyr Trp Tyr Cys Thr Thr Ala Pro Tyr Asn Val
1               5                  10                  15
```

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(i x) FEATURE:
(A) NAME/KEY: Region
(B) LOCATION: 1..16
(D) OTHER INFORMATION: /label=VH

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..16
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:74:

```
Thr Glu Ile Glu Trp Tyr Gly Cys Asn Leu Arg Thr Thr Phe Thr Arg
1               5                  10                  15
```

(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

-continued

(i i) MOLECULE TYPE: protein

(i x) FEATURE:
(A) NAME/KEY: Region
(B) LOCATION: 1..22
(D) OTHER INFORMATION: /label=VH

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..22
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:75:

```
Asn Gln Leu Ala Gly Gly Trp Tyr Leu Asp Pro Asn Tyr Trp Leu Ser
1               5                  10                  15

Val Gly Ala Tyr Ala Ile
            20
```

(2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(i x) FEATURE:
(A) NAME/KEY: Region
(B) LOCATION: 1..24
(D) OTHER INFORMATION: /label=VH

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..24
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:76:

```
Arg Leu Thr Glu Met Gly Ala Cys Asp Ala Arg Trp Ala Thr Leu Ala
1               5                  10                  15

Thr Arg Thr Phe Ala Tyr Asn Tyr
            20
```

(2) INFORMATION FOR SEQ ID NO:77:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(i x) FEATURE:
(A) NAME/KEY: Region
(B) LOCATION: 1..24
(D) OTHER INFORMATION: /label=VH

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..24
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:77:

```
Asp Gly Trp Thr Arg Lys Glu Gly Gly Ile Gly Leu Pro Trp Ser Val
1               5                  10                  15

Gln Cys Glu Asp Gly Tyr Asn Tyr
            20
```

(2) INFORMATION FOR SEQ ID NO:78:

-continued

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(i x) FEATURE:
(A) NAME/KEY: Region
(B) LOCATION: 1..10
(D) OTHER INFORMATION: /label=VH

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..10
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:78:

Asp Ser Tyr Pro Cys His Leu Leu Asp Val
1 5 10

(2) INFORMATION FOR SEQ ID NO:79:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(i x) FEATURE:
(A) NAME/KEY: Region
(B) LOCATION: 1..12
(D) OTHER INFORMATION: /label=VH

(i x) FEATURE:
(A) NAME/KEY: Domain
(B) LOCATION: 1..12
(D) OTHER INFORMATION: /label=CDR3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:79:

Val Glu Tyr Pro Ile Ala Asp Met Cys Ser Arg Tyr
1 5 10

(2) INFORMATION FOR SEQ ID NO:80:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(v i) ORIGINAL SOURCE:
(A) ORGANISM: Camelus dromedarius

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:80:

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly
20 25

(2) INFORMATION FOR SEQ ID NO:81:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

-continued

(i i) MOLECULE TYPE: protein

(v i) ORIGINAL SOURCE:
(A) ORGANISM: Camelus dromedarius

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:81:

```
Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val Ser
1               5                  10
```

(2) INFORMATION FOR SEQ ID NO:82:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:82:

```
Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr Leu Gln
1               5                  10                  15

Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ala Arg
            20                  25                  30
```

(2) INFORMATION FOR SEQ ID NO:83:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:83:

```
Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Gly Thr Asn Glu Val
1               5                  10                  15

Cys Lys Cys Pro Lys Cys Pro Ala Pro Glu Leu Pro Gly Gly Pro Ser
            20                  25                  30

Val Phe Val Phe Pro
        35
```

(2) INFORMATION FOR SEQ ID NO:84:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:84:

```
Gly Gly Ser Val Gln Gly Gly Gly Ser Leu Arg Leu Ser Cys Ala Ile
1               5                  10                  15

Ser Gly
```

(2) INFORMATION FOR SEQ ID NO:85:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

-continued

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:85:

```
Trp Phe Arg Glu Gly Pro Gly Lys Glu Arg Glu Gly Ile Ala
1               5                  10
```

(2) INFORMATION FOR SEQ ID NO:86:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:86:

```
Arg Phe Thr Ile Ser Gln Asp Ser Thr Leu Lys Thr Met Tyr Leu Leu
1               5                  10                  15

Met Asn Asn Leu Lys Pro Glu Asp Thr Gly Thr Tyr Tyr Cys Ala Ala
            20                  25                  30
```

(2) INFORMATION FOR SEQ ID NO:87:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 60 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:87:

```
Trp Gly Gln Gly Thr Gln Val Thr Val Ser Ser Glu Pro Lys Ile Pro
1               5                  10                  15

Gln Pro Gln Pro Lys Pro Gln Pro Gln Pro Gln Pro Gln Pro Lys Pro
            20                  25                  30

Gln Pro Lys Pro Glu Pro Glu Cys Thr Cys Pro Lys Cys Pro Ala Pro
        35                  40                  45

Glu Leu Leu Gly Gly Pro Ser Val Phe Ile Phe Pro
    50                  55                  60
```

(2) INFORMATION FOR SEQ ID NO:88:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:88:

```
Gly Gly Ser Val Gln Ala Gly Gly Ser Leu Arg Leu Ser Cys Ala Ser
1               5                  10                  15

Ser Ser
```

(2) INFORMATION FOR SEQ ID NO:89:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:89:

-continued

```
Trp Tyr Arg Gln Ala Pro Gly Lys Glu Arg Glu Phe Val Ser
1               5                  10
```

(2) INFORMATION FOR SEQ ID NO:90:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:90:

```
Arg Phe Thr Ile Ser Gln Asp Ser Ala Lys Asn Thr Val Tyr Leu Gln
1               5                  10                  15

Met Asn Ser Leu Lys Pro Glu Asp Thr Ala Met Tyr Tyr Cys Lys Ile
            20                  25                  30
```

(2) INFORMATION FOR SEQ ID NO:91:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:91:

```
Trp Gly Gln Gly Thr Gln Val Thr Val Ser Ser Gly Thr Asn Glu Val
1               5                  10                  15

Cys Lys Cys Pro Lys Cys Pro Ala Pro Glu Leu Pro Gly Gly Pro Ser
            20                  25                  30

Val Phe Val Phe Pro
        35
```

(2) INFORMATION FOR SEQ ID NO:92:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 400 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: cDNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:92:

```
CTCGAGTCTG GGGGAGGATC GGTGCAGGCT GGAGGGTCTC TGAGACTCTC GTGCGCAGCC      60

TCTGGATACA GTAATTGTCC CCTCACTTGG AGCTGGTATC GCCAGTTTCC AGGAACGGAG     120

CGCGAGTTCG TCTCCAGTAT GGATCCGGAT GGAAATACCA AGTACACATA CTCCGTGAAG     180

GGCCGCTTCA CCATGTCCCG AGGCAGCACC GAGTACACAG TATTTCTGCA AATGGACAAT     240

CTGAAACCTG AGGACACGGC GATGTATTAC TGTAAAACAG CCCTACAACC TGGGGGTTAT     300

TGTGGGTATG GGTANTGCCT CTGGGGCCAG GGGACCCAGG TCACCGTCTC CTCACTAGTT     360

ACCCGTACGA CGTTCCGGAC TACGGTTCTT AATAGAATTC                           400
```

(2) INFORMATION FOR SEQ ID NO:93:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 391 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

-continued

(i i) MOLECULE TYPE: cDNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:93:

```
CTCGAGTCTG GGGGAGGCTC GGTGCAGGCT GGAGGGTCTC TGAGACTCTC CTGTGCATCT      60
TCTTCTAAAT ATATGCCTTG CACCTACGAC ATGACCTGGT ACCGCCAGGC TCCAGGCAAG     120
GAGCGCGAAT TTGTCTCAAG TATAAATATT GATGGTAAGA CAACATACGC AGACTCCGTG     180
AAGGGCCGAT TCACCATCTC CCAAGACAGC GCCAAGAACA CGGTGTATCT GCAGATGAAC     240
AGCCTGAAAC CTGAGGACAC GGCGATGTAT TACTGTAAAA TAGATTCGTA CCCGTGCCAT     300
CTCCTTGATG TCTGGGGCCA GGGGACCCAG GTCACCGTCT CCTCACTAGT TACCCGTACG     360
AGCTTCCGGA CTACGGTTCT TAATAGAATT C                                    391
```

(2) INFORMATION FOR SEQ ID NO:94:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 443 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: cDNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:94:

```
CAGGTGAAAC TGCTCGAGTC TGGAGGAGGC TCGGTGCAGA CTGGAGGATC TCTGAGACTC      60
TCCTGTGCAG TCTCTGGATT CTCCTTTAGT ACCAGTTGTA TGGCCTGGTT CCGCCAGGCT     120
TCAGGAAAGC AGCGTGAGGG GGTCGCAGCC ATTAATAGTG GCGGTGGTAG GACATACTAC     180
AACACATATG TCGCCGAGTC CGTGAAGGGC CGATTCGCCA TCTCCCAAGA CAACGCCAAG     240
ACCACGGTAT ATCTTGATAT GAACAACCTA ACCCCTGAAG ACACGGCTAC GTATTACTGT     300
GCGGCGGTCC CAGCCCACTT GGGACCTGGC GCCATTCTTG ATTTGAAAAA GTATAAGTAC     360
TGGGGCCAGG GGACCCAGGT CACCGTCTCC TCACTAGCTA GTTACCCGTA CGACGTTCCG     420
GACTACGGTT CTTAATAGAA TTC                                             443
```

(2) INFORMATION FOR SEQ ID NO:95:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 433 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: cDNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:95:

```
CTCGAGTCTG GGGGAGGGTC GGTGCAGGCT GGAGGGTCTC TGAGACTCTC CTGTAATGTC      60
TCTGGCTCTC CCAGTAGTAC TTATTGCCTG GGCTGGTTCC GCCAGGCTCC AGGGAGGGAG     120
CGTGAGGGGG TCACAGCGAT TAACACTGAT GGCAGTATCA TATACGCAGC CGACTCCGTG     180
AAGGGCCGAT TCACCATCTC CCAAGACACC GCCAAGGAAA CGGTACATCT CCAGATGAAC     240
AACCTGCAAC CTGAGGATAC GGCCACCTAT TACTGCGCGG CAAGACTGAC GGAGATGGGG     300
GCTTGTGATG CGAGATGGGC GACCTTAGCG ACAAGGACGT TTGCGTATAA CTACTGGGGC     360
CGGGGGACCC AGGTCACCGT CTCCTCACTA GTTACCCGTA CGACGTTCCG GACTACGGTT     420
CTTAATAGAA TTC                                                        433
```

(2) INFORMATION FOR SEQ ID NO:96:

-continued

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 449 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: cDNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:96:

```
CAGGTGAAAC TGCTCGAGTC TGGGGGAGGG TCGGTGCAGG CTGGAGGGTC TCTGAGACTC      60
TCCTGTAATG TCTCTGGCTC TCCCAGTAGT ACTTATTGCC TGGGCTGGTT CCGCCAGGCT     120
CCAGGGAAGG AGCGTGAGGG GGTCACAGCG ATTAACACTG ATGGCAGTGT CATATACGCA     180
GCCGACTCCG TGAAGGGCCG ATTCACCATC TCCCAAGACA CCGCCAAGAA AACGGTATAT     240
CTCCAGATGA ACAACCTGCA ACCTGAGGAT ACGGCCACCT ATTACTGCGC GGCAAGACTG     300
ACGGAGATGG GGGCTTGTGA TGCGAGATGG GCGACCTTAG CGACAAGGAC GTTTGCGTAT     360
AACTACTGGG GCCGGGGGAC CCAGGTCACC GTCTCCTCAC TAGCTAGTTA CCCGTACGAC     420
GTTCCGGACT ACGGTTCTTA ATAGAATTC                                       449
```

(2) INFORMATION FOR SEQ ID NO:97:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 424 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: cDNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:97:

```
CTCGAGTCTG GAGGAGGCTC GGCGCAGGCT GGAGGATCTC TGAGACTCTC CTGTGCAGCC      60
CACGGGATTC CGCTCAATGG TTACTACATC GCCTGGTTCC GTCAGGCTCC TGGGAAGGGG     120
CGTGAGGGGG TCGCAACAAT TAATGGTGGT CGCGACGTCA CATACTACGC CGACTCCGTG     180
ACGGGCCGAT TTACCATCTC CGAGACAGC CCCAAGAATA CGGTGTATCT GCAGATGAAC      240
AGCCTGAAAC CTGAGGACAC GGCCATCTAC TTCTGTGCAG CAGGCTCGCG TTTTTCTAGT     300
CCTGTTGGGA GCACTTCTAG ACTCGAAAGT AGCGACTATA ACTATTGGGG CCAGGGGATC     360
CAGGTCACCG TCACCTCACT AGTTACCCGT ACGACGTTCC GGACTACGGT TCTTAATAGA     420
ATTC                                                                 424
```

(2) INFORMATION FOR SEQ ID NO:98:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 415 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: cDNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:98:

```
CTCGAGTCTG GAGGAGGCTC GGTTCAGGCT GGAGGGTCCC TTAGACTCTC CTGTGCAGCC      60
TCTGACTACA CCATCACTGA TTATTGCATG GCCTGGTTCC GCCAGGCTCC AGGGAAGGAG     120
CGTGAATTGG TCGCAGCGAT TCAAGTTGTC CGTAGTGATA CTCGCCTCAC AGACTACGCC     180
GACTCCGTGA AGGGACGATT CACCATCTCC CAAGGCAACA CCAAGAACAC AGTGAATCTG     240
CAAATGAACA GCCTGACACC TGAGGACACG GCCATCTACA GTTGTGCGGC AACCAGTAGT     300
TTTTACTGGT ACTGCACCAC GGCGCCTTAT AACGTCTGGG GTCAGGGGAC CCAGGTCACC     360
```

-continued

GTCTCCTCAC TAGTTACCCG TACGACGTTC CGGACTACGG TTCTTAATAG AATTC 415

(2) INFORMATION FOR SEQ ID NO:99:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 406 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: cDNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:99:

CTCGAGTCTG GGGGAGGCTC GGTGCAGGGT GGAGGGTCTC TGAGACTCTC CTGTGCAATC 60

TCTGGATACA CGTACGGTAG CTTCTGTATG GGCTGGTTCC GCGAGGGTCC AGGCAAGGAA 120

CGTGAGGGGA TCGCAACTAT TCTTAATGGT GGTACTAACA CATACTATGC CGACTCGGTG 180

AAGGGCCGAT TCACCATCTC CCAAGACAGC ACGTTGAAGA CGATGTATCT GCTAATGAAC 240

AACCTGAAAC CTGAAGACAC GGGCACCTAT TACTGTGCTG CAGAACTAAG TGGTGGTAGT 300

TGTGAATTGC CTTTGCTATT TGACTACTGG GGCCAGGGCA CCCAGGTCAC CGTCTCCTCA 360

CTAGTTACCC GTACGACGTT CCGGACTACG GTTCTTAATA GAATTC 406

(2) INFORMATION FOR SEQ ID NO:100:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 427 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: cDNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:100:

CTCGAGTCTG GGGGAGGCTC GGTGCAGGCT GGAGGGTCTC TGAGACTCTC CTGTACAGGC 60

TCTGGATTCC CCTATAGTAC CTTCTGTCTG GGGTGGTTCC GCCAGGCTCC AGGGAAGGAG 120

CGTGAGGGGG TCGCGGGTAT TAATAGTGCA GGAGGTAATA CTTACTATGC CGACGCCGTG 180

AAGGGCCGAT TCACCATCTC CCAAGGGAAT GCCAAGAATA CGGTGTTTCT GCAAATGGAT 240

AACTTGAAAC CTGAGGACAC GGCCATCTAT TACTGCGCGG CGGATAGTCC ATGTTACATG 300

CCGACTATGC CCGCTCCCCC GATACGAGAC AGTTTTGGCT GGGATGATTT TGGCCAGGGG 360

ACCCAGGTCA CCGTCTCCTC ACTAGTTACC CGTACGACGT TCCGGACTAC GGTTCTTAAT 420

AGAATTC 427

(2) INFORMATION FOR SEQ ID NO:101:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 409 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: cDNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:101:

CTCGAGTCAG GGGGAGGCTC GGTACAGGTT GGAGGGTCTC TGAGACTCTC CTGTGTAGCC 60

TCTACTCACA CCGACAGTAG CACCTGTATA GGCTGGTTCC GCCAGGCTCC AGGGAAGGAG 120

CGCGAGGGGG TCGCAAGTAT ATATTTTGGT GATGGTGGTA CGAATTATCG CGACTCCGTG 180

AAGGGCCGAT TCACCATCTC CCAACTCAAC GCCCAGAACA CAGTGTATCT GCAAATGAAC 240

AGCCTGAAAC CTGAGGACAG CGCCATGTAC TACTGTGCAA TCACTGAAAT TGAGTGGTAT 300

-continued

GGGTGCAATT TAAGGACTAC TTTTACTCGC TGGGGCCAGG GGACCCAGGT CACCGTCTCC 360

TCACTAGTTA CCCGTACGAC GTTCCGGACT ACGGTTCTTA ATAGAATTC 409

(2) INFORMATION FOR SEQ ID NO:102:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 445 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: cDNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:102:

CTCGAGTCTG GGGGAGGCTC GGTACAAACT GGAGGGTCTC TGAGACTCTC TTGCGAAATC 60

TCTGGATTGA CTTTTGATGA TTCTGACGTG GGGTGGTACC GCCAGGCTCC AGGGGATGAG 120

TGCAAATTGG TCTCAGGTAT TCTGAGTGAT GGTACTCCAT ATACAAAGAG TGGAGACTAT 180

GCTGAGTCTG TGAGGGGCCG GGTTACCATC TCCAGAGACA ACGCCAAGAA CATGATATAC 240

CTTCAAATGA ACGACCTGAA ACCTGAGGAC ACGGCCATGT ATTACTGCGC GGTAGATGGT 300

TGGACCCGGA AGGAAGGGGG AATCGGGTTA CCCTGGTCGG TCCAATGTGA AGATGGTTAT 360

AACTATTGGG GCCAGGGGAC CCAGGTCACC GTCTCCTCAC TAGTTACCCG TACGACGTTC 420

CGGACTACGG TTCTTAATAG AATTC 445

(2) INFORMATION FOR SEQ ID NO:103:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 394 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: cDNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:103:

CTCGAGTCTG GAGGAGGCTC GGTGCAGGCT GGAGGGTCTC TGAGACTCTC CTGTGTAGCC 60

TCTGGATTCA ATTTCGAAAC TTCTCGTATG GCGTGGTACC GCCAGACTCC AGGAAATGTG 120

TGTGAGTTGG TCTCAAGTAT TTACAGTGAT GGCAAAACAT ACTACGTCGA CCGCATGAAG 180

GGCCGATTCA CCATTTCTAG AGAGAATGCC AAGAATACAT TGTATCTACA ACTGAGCGGC 240

CTCAAACCTG AGGACACGGC CATGTATTAC TGTGCGCCGG TTGAATATCC TATTGCAGAC 300

ATGTGTTCGA GATACGGCGA CCCGGGGACC CAGGTCACCG TCTCCTCACT AGTTACCCGT 360

ACGACGAACC GGACTACGGT TCTTAATAGA ATTC 394

(2) INFORMATION FOR SEQ ID NO:104:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 433 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: cDNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:104:

CTCGAGTCTG GGGGAGGCTC GGTGCAGGTT GGAGGGTCTC TGAAACTCTC CTGTAAAATC 60

TCTGGAGGTA CCCCAGATCG TGTTCCTAAA TCTTTGGCCT GGTTCCGCCA GGCTCCAGAG 120

AAGGAGCGCG AGGGGATCGC AGTTCTTTCG ACTAAGGATG GTAAGACATT CTATGCCGAC 180

-continued

```
TCCGTGAAGG GCCGATTCAC CATCTTCTTA GATAATGACA AGACCACTTT CTCCTTACAA   240
CTTGATCGAC TGAACCCGGA GGACACTGCC GACTACTACT GCGCTGCAAA TCAATTAGCT   300
GGTGGCTGGT ATTTGGACCC GAATTACTGG CTCTCTGTGG GTGCATATGC CATCTGGGGC   360
CAGGGGACCC AGGTCACCGT CTCCTCACTA GTTACCCGTA CGACGTTCCG GACTACGGTT   420
CTTAATAGAA TTC                                                      433
```

(2) INFORMATION FOR SEQ ID NO:105:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 416 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: cDNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:105:

```
CAGGTGAAAC TGCTCGAGTC TGGGGGAGGC TCGGTGCAGG CTGGGGGGTC TCTGACACTC    60
TCTTGTGTAT ACACCAACGA TACTGGGACC ATGGGATGGT TTCGCCAGGC TCCAGGGAAA   120
GAGTGCGAAA GGGTCGCGCA TATTACGCCT GATGGTATGA CCTTCATTGA TGAACCCGTG   180
AAGGGGCGAT TCACGATCTC CCGAGACAAC GCCCAGAAAA CGTTGTCTTT GCGAATGAAT   240
AGTCTGAGGC CTGAGGACAC GGCCGTGTAT TACTGTGCGG CAGATTGGAA ATACTGGACT   300
TGTGGTGCCC AGACTGGAGG ATACTTCGGA CAGTGGGGTC AGGGGGCCCA GGTCACCGTC   360
TCCTCACTAG CTAGTTACCC GTACGACGTT CCGGACTACG GTTCTTAATA GAATTC       416
```

(2) INFORMATION FOR SEQ ID NO:106:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 361 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: cDNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:106:

```
CTCGAGTCTG GGGGAGGCTC GGTCCAACCT GGAGGATCTC TGACACTCTC CTGTACAGTT    60
TCTGGGGCCA CCTACAGTGA CTACAGTATT GGATGGATCC GCCAGGCTCC AGGGAAGGAC   120
CGTGAAGTAG TCGCAGCCGC TAATACTGGT GCGACTAGTA AATTCTACGT CGACTTTGTG   180
AAGGGCCGAT TCACCATTTC CCAAGACAAC GCCAAGAATA CGGTATATCT GCAAATGAGC   240
TTCCTGAAAC CTGAGGACAC GGCCATCTAT TACTGTGCGG CAGCGGACCC AAGTATATAT   300
TATAGTATCC TCCATTGAGT ATAAGTACTG GGGCCAGGGG ACCCAGGTCA CCGTCTCCTC   360
A                                                                   361
```

(2) INFORMATION FOR SEQ ID NO:107:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 354 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: cDNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:107:

```
CTCGAGTCAG GGGGAGGCTC GGTGGAGGCT GGAGGGTCTC TGAGACTCTC CTGTACAGCC    60
TCTGGATACG TATCCTCTAT GGCCTGGTTC CGCCAGGTTC CAGGGCAGGA GCGCGAGGGG   120
```

-continued

GTCGCGTTTG TTCAAACGGC TGACAATAGT GCATTATATG GCGACTCCGT GAAGGGCCGA 180

TTCACCATCT CCCACGACAA CGCCAAGAAC ACGCTGTATC TGCAAATGCG CAACCTGCAA 240

CCTGACGACA CTGGCGTGTA CTACTGTGCG GCCCAAAAGA AGGATCGTAC TAGATGGGCC 300

GAGCCTCGAG AATGGAACAA CTGGGGCCAG GGGACCCAGG TCACCGTCTC CTCA 354

(2) INFORMATION FOR SEQ ID NO:108:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 381 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: cDNA

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:108:

CTCGAGTCAG GTGTCCGGTC TGATGTGCAG CTGGTGGCGT CTGGGGGAGG CTCGGTGCAG 60

GCTGGAGGCT CTCTGAGACT CTCCTGTACA GCCTCTGGAG ACAGTTTCAG TAGATTTGCC 120

ATGTCTTGGT TCCGCCAGGC TCCAGGGAAG GAGTGCGAAT TGGTCTCAAG CATTCAAAGT 180

AATGGAAGGA CAACTGAGGC CGATTCCGTG CAAGGCCGAT TCACCATCTC CCGAGACAAT 240

TCCAGGAACA CAGTGTATCT GCAAATGAAC AGCCTGAAAC CCGAGGACAC GGCCGTGTAT 300

TACTGTGGGG CAGTCTCCCT AATGGACCGA ATTTCCCAAC ATGGGTGCCG GGGCCAGGGA 360

ACCCAGGTCA CCGTCTCCTT A 381

(2) INFORMATION FOR SEQ ID NO:109:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:109:

Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp
1 5 10 15

Glu Leu

(2) INFORMATION FOR SEQ ID NO:110:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:110:

Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu
1 5 10 15

Glu Met

(2) INFORMATION FOR SEQ ID NO:111:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

-continued

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:111:

Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu
1 5 10 15

Glu Met

(2) INFORMATION FOR SEQ ID NO:112:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:112:

Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp
1 5 10 15

Thr Leu Met Ile Ser Arg Thr Pro
20

(2) INFORMATION FOR SEQ ID NO:113:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:113:

Val Ala Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr
1 5 10 15

Leu Met Ile Ser Arg Thr Pro
20

(2) INFORMATION FOR SEQ ID NO:114:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:114:

Phe Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp
1 5 10 15

Thr Leu Met Ile Ser Arg Thr Pro
20

(2) INFORMATION FOR SEQ ID NO:115:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:115:

Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg
1 5 10 15

Asp Glu Leu

(2) INFORMATION FOR SEQ ID NO:116:

-continued

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:116:

Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg
1 5 10 15

Glu Glu Met

(2) INFORMATION FOR SEQ ID NO:117:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:117:

Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Gln
1 5 10 15

Glu Glu Met

(2) INFORMATION FOR SEQ ID NO:118:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:118:

Glu Val Lys Leu Val Glu Ser Gly Gly Gly Leu Val Glu Pro Gly Gly
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Thr Ser Gly Phe Thr Phe Ser
20 25 30

(2) INFORMATION FOR SEQ ID NO:119:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:119:

Glu Val Gln Leu Leu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly Ser
1 5 10 15

Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser
20 25

(2) INFORMATION FOR SEQ ID NO:120:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:120:

-continued

```
Trp Gly Gln Gly Thr Thr Val Thr Val Ser Ser
1               5                  10
```

(2) INFORMATION FOR SEQ ID NO:121:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:121:

```
Trp Gly Gln Gly Thr Met Val Thr Val Ser Ser
1               5                  10
```

(2) INFORMATION FOR SEQ ID NO:122:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:122:

```
Trp Gly Gln Gly Thr Thr Leu Thr Val Ser Ser
1               5                  10
```

(2) INFORMATION FOR SEQ ID NO:123:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:123:

```
Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ala
1               5                  10
```

(2) INFORMATION FOR SEQ ID NO:124:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:124:

```
Trp Gly Ala Gly Thr Thr Val Thr Val Ser Ser
1               5                  10
```

(2) INFORMATION FOR SEQ ID NO:125:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:125:

```
Asp Tyr Tyr Gly Ser Ser Tyr Phe Asp Val Trp Gly Ala Gly Thr Thr
1               5                  10                  15

Val Thr Val Ser Ser
            20
```

-continued

(2) INFORMATION FOR SEQ ID NO:126:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 67 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:126:

Lys Val Asp Lys Arg Val Glu Leu Lys Thr Pro Leu Gly Asp Thr Thr
1 5 10 15

His Thr Cys Pro Arg Cys Pro Glu Pro Lys Cys Ser Asp Thr Pro Pro
20 25 30

Pro Cys Pro Arg Cys Pro Glu Pro Lys Ser Cys Asp Thr Pro Pro Pro
35 40 45

Cys Pro Arg Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe
50 55 60

Leu Phe Pro
65

(2) INFORMATION FOR SEQ ID NO:127:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:127:

Lys Val Asp Lys Lys Ala Glu Pro Lys Ser Cys Asp Lys Thr His Thr
1 5 10 15

Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe
20 25 30

Leu Phe Pro
35

(2) INFORMATION FOR SEQ ID NO:128:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:128:

Lys Val Lys Val Thr Val Glu Arg Lys Cys Cys Val Glu Cys Pro Pro
1 5 10 15

Cys Pro Ala Pro Pro Val Ala Gly Pro Ser Val Phe Leu Phe Pro
20 25 30

(2) INFORMATION FOR SEQ ID NO:129:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:129:

Lys Val Asp Lys Arg Val Glu Ser Lys Tyr Gly Pro Pro Cys Pro Ser

-continued

1	5	10	15
Cys Pro Ala Pro Glu Phe Leu Gly Gly Pro Ser Val Phe Leu Phe Pro			
20		25	30

(2) INFORMATION FOR SEQ ID NO:130:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:130:

1	5	10
Trp Gly Arg Gly Thr Leu Val Thr Val Ser Ser		

We claim:

1. Nucleotide sequence encoding an immunoglobulin or a fragment of an immunoglobulin selected from the group consisting of the V_{HH}, hinge, $FV_{HH}h$, $F(V_{HH}h)_2$, CDR, CDR_1, CDR_3, FW, C_H2, and C_H3, said immunoglobulin comprising two heavy polypeptide chains, each heavy chain consisting of a complete antigen binding site, said immunoglobulin containing a variable (V_{HH}) region and a constant region, said constant region being devoid of first constant domain C_H1, wherein the immunoglobulin is devoid of polypeptide light chains, which immunoglobulin comprises a peptide sequence selected from the group consisting of:

VTVSSGTNEVCKCPKCPAPELPGGPSVVFVVFP (SEQ ID NO:43),
VTVSSEPKIPQPQPKPQPQPQPQPKPQPQPEPE
CTCPKCPAPELLGGPSVFIFP (SEQ ID NO:44)
GTNEVCKCPKCP (SEQ ID NO:37)
APELPGGPSVFVFP (SEQ ID NO:45)
EPKIPQPQPKPQPQPQPQPQPPKPQPKPEPEECTCPKCP (SEQ ID NO:38)
APELLGGPSVFIFP (SEQ ID NO:46)
APELLGGPTVFIFPPKPKDVLSITLTP (SEQ ID NO:31)
APELPGGPSVFVFPTKPKDVLSISGRP (SEQ ID NO:32)
APELPGGPSVFVFPPKPKDVLSISGRP (SEQ ID NO:33)
APELLGGPSVFIFPPKPKDVLSISGRP (SEQ ID NO:34)
GQTREPQVYTLA (SEQ ID NO:35)
GQTREPQVYTLAPXRLEL (SEQ ID NO:36)
GQPREPQVYTLPPSRDEL (SEQ ID NO:109)
GQPREPQVYTLPPSREEM (SEQ ID NO:110)
GQPREPQVYTLPPSQEEM (SEQ ID NO:111)
GGSVQTGGSLRLSCEISGLTFD (SEQ ID NO:1)
GGSVQTGGSLRLSCAVSGFSFS (SEQ ID NO:2)
GGSEQGGGSLRLSCAISGYTYG (SEQ ID NO:3)
GGSVQPGGSLTLSCTVSGATYS (SEQ ID NO:4)
GGSVQAGGSLRLSCTGSGFPYS (SEQ ID NO:5)
GGSVQAGGSLRLSCVAGFGTS (SEQ ID NO:6)
GGSVQAGGSLRLSCVSFSPSS (SEQ ID NO:7)
WGQGTQVTVSS (SEQ ID NO:8)
WGQGTLVTVSS (SEQ ID NO:9)
WGQGAQVTVSS (SEQ ID NO:10)
WGQGTQVTASS (SEQ ID NO:11)
RGQGTQVTVSL (SEQ ID NO:12)
and/or,
ALQPGGYCGYGX----------CL (SEQ ID NO:62)
VSLMDRISQH----------GC (SEQ ID NO:63)
VPAHLGPGAILDLKKY------KY (SEQ ID NO:64)
FCYSTAGDGGSGE---------MY (SEQ ID NO:65)
ELSGGSCELPLLF---------DY (SEQ ID NO:66)
DWKYWTCGAQTGGYF-------GQ (SEQ ID NO:67)
RLTEMGACDARWATLATRTFAYNY (SEQ ID NO:68)
QKKDRTRWAEPREW--------NN (SEQ ID NO:69)
GSRFSSPVGSTSRLES-SDY--NY (SEQ ID NO:70)
ADPSIYYSILXIEY--------KY (SEQ ID NO:71)
DSPCYMPTMPAPPIRDSFGW--DD (SEQ ID NO:72)

-continued

TSSFYWYCTTAPY---------NV (SEQ ID NO:73)
TEIEWYGCNLRTTF--------TR (SEQ ID NO:74)
NQLAGGWYLDPNYWLSVGAY--AI (SEQ ID NO:75)
RLTEMGACDARWATLATRTFAYNY (SEQ ID NO:76)
DGWTRKEGGIGLPWSVQCEDGYNY (SEQ ID NO:77)
DSYPCHLL--------------DV (SEQ ID NO:78);
and
VEYPIADMCS-----------RY (SEQ ID NO:79).

2. Process for the preparation of antibodies directed against determined antigens, comprising the steps of:

cloning into vectors, especially into phages and more particularly filamentous bacteriophages, a DNA or cDNA sequence obtained from lymphocytes of Camelids previously immunized with determined antigens, capable of producing an immunoglobulin comprising two heavy polypeptide chains, each heavy chain consisting of a complete antigen binding site, said immunoglobulin containing a variable (V_{HH}) region and a constant region, said constant region being devoid of first constant domain C_H1, wherein the immunoglobulin is devoid of polypeptide light chains;

transforming prokaryotic cells with said vectors in conditions allowing the production of the antibodies;

selecting the appropriate antibody by subjecting the transforming cells to antigen-affinity selection; and recovering the antibodies having the desired specificity.

3. Recombinant vector comprising a nucleotide sequence according to claim 1, wherein the vector is a plasmid, a phage especially a bacteriophage, a virus, a YAC, or a cosmid.

4. Recombinant cell or organism modified by a vector as claimed in claim 3.

5. An isolated nucleic acid encoding an immunoglobulin comprising two heavy polypeptide chains, each heavy chain consisting of a complete antigen binding site, said immunoglobulin containing a variable (V_{HH}) region and a constant region, said constant region being devoid of first constant domain C_H1, wherein the immunoglobulin is devoid of polypeptide light chains, wherein said nucleic acid comprises a nucleic acid fragment having a sequence selected from SEQ ID NOS: 92–108.

6. A process for the preparation of antibodies directed against determined antigens, comprising the steps of:

cloning into a vector, a DNA or cDNA sequence obtained from lymphocytes of Camelids previously immunized with determined antigens,

transforming prokaryotic cells with said vector in conditions allowing the production of the antibodies,

selecting antibody by subjecting the transforming cells to antigen-affinity selection, and

recovering the antibodies having the desired specificity.

7. The process according to claim 6, wherein the cloning vector is a plasmid or a eukaryotic virus and the transformed cell is a eukaryotic cell.

8. The process of claim 7 wherein said eukaryotic cell is selected from the group consisting of a yeast cell, mammalian cell, plant cell and protozoan cell.

* * * * *